

Der Rolling Cross-Section Survey - ein Instrument zur Analyse dynamischer Prozesse der Einstellungsentwicklung: Bericht zur ersten deutschen RCS-Studie anlässlich der Bundestagswahl 2005

Schmitt-Beck, Rüdiger; Faas, Thorsten; Holst, Christian

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Zur Verfügung gestellt in Kooperation mit / provided in cooperation with:
GESIS - Leibniz-Institut für Sozialwissenschaften

Empfohlene Zitierung / Suggested Citation:

Schmitt-Beck, R., Faas, T., & Holst, C. (2006). Der Rolling Cross-Section Survey - ein Instrument zur Analyse dynamischer Prozesse der Einstellungsentwicklung: Bericht zur ersten deutschen RCS-Studie anlässlich der Bundestagswahl 2005. *ZUMA Nachrichten*, 30(58), 13-49. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-211128>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

DER ROLLING CROSS-SECTION SURVEY – EIN INSTRUMENT ZUR ANALYSE DYNAMISCHER PROZESSE DER EINSTELLUNGSENTWICKLUNG

Bericht zur ersten deutschen RCS-Studie anlässlich der
Bundestagswahl 2005

THE ROLLING CROSS-SECTION SURVEY – AN INSTRUMENT FOR ANALYZING THE DYNAMICS OF PUBLIC OPINION

A Documentation of the First German RCS Survey Conducted in the
Context of the 2005 National Election

RÜDIGER SCHMITT-BECK, THORSTEN FAAS & CHRISTIAN HOLST

Der Beitrag stellt eine neue, Panel-Befragungen komplementäre Methode zur Analyse dynamischer Prozesse der Meinungskristallisation und des Meinungswandels vor: den *Rolling Cross-Section Survey*. Die Grundidee des RCS-Designs besteht darin, die Durchführung einer auf einer Zufallsstichprobe basierenden Querschnittsbefragung in strikt kontrollierter Weise so über einen vorab definierten Zeitraum zu spreizen, dass die Befragten jedes einzelnen Tages jeweils in sich eine Zufallsstichprobe aus der Grundgesamtheit konstituieren. Solche Umfragen ermöglichen in höchst flexibler Weise Analysen von dynamischen Phänomenen der öffentlichen Meinung. Der Beitrag erörtert detailliert die idealtypischen Eigenschaften von RCS-Studien und diskutiert Probleme ihrer praktischen Umsetzung am Beispiel der ersten deutschen RCS-Studie, die während des Wahlkampfes zur Bundestagswahl 2005 realisiert wurde. Er schließt mit einigen exemplarischen Analysen, welche Rückschlüsse auf die Datenqualität ermöglichen und die besonderen Vorzüge von RCS-Daten verdeutlichen.

The present paper describes a new methodological approach for analysing dynamic processes of opinion crystallization and change: the *Rolling Cross-Section Survey*. The basic idea of the RCS design is to spread a regular cross-section survey that is based on a random sample over an a priori defined period of time in such a way that not only the entire sample, but also the respondents interviewed on any given day of the field period

constitute random samples. Data generated from RCS surveys give researchers unique opportunities to study dynamic phenomena of public opinion in highly flexible ways. Following an overview of the ideal-typical features of the RCS design, the paper discusses problems of implementing these principles in practical fieldwork, taking the example of the first German RCS study that was conducted during the campaign of the German National Parliamentary Election 2005. The paper closes with exemplary analyses that indicate the quality of these data and illustrate their usefulness.

1 Einleitung

Bevölkerungsumfragen haben in den letzten Jahrzehnten sowohl in der politikwissenschaftlichen Forschung als auch in der medialen Beobachtung des politischen Prozesses weite Verbreitung gefunden. In beiden Bereichen stellen sie heute so etwas dar wie den „Goldstandard“ (Brady 2000: 47) für die Messung der Orientierungen der Bürger zum politischen System, zu politischen Akteuren, politischen Problemen und den diskutierten Lösungsvorschlägen. Aus politikwissenschaftlicher Perspektive bilden Umfragen ein ausgezeichnetes Instrument, um zu testen, wie sich Theorien in der Konfrontation mit empirischen Beobachtungen bewähren. Gerade das Ziel, Theorien mittels Umfragen zu prüfen, war immer wieder ein Ansporn für methodische Innovationen. Computer-unterstützte Telefoninterviews, aber auch Möglichkeiten der Internet-basierten Befragung stellen nur zwei Beispiele dar, wie technische Neuerungen neue Möglichkeiten für die Umfrageforschung eröffnen.

Eine dieser Innovationen der Umfrageforschung steht im Mittelpunkt des vorliegenden Beitrages: das so genannte „*Rolling Cross-Section*“-Design (im Folgenden: RCS). Das zentrale Ziel von RCS-Studien besteht darin, dynamische Prozesse der Kristallisierung und Veränderung von Einstellungen und anderen Orientierungen abzubilden und zu analysieren. Anwendungen finden sich bisher ausschließlich im Bereich der Wahlkampfforschung. Im Prinzip stellt das RCS-Design aber auch für andere Phänomene kurzfristigen Wandels der öffentlichen Meinung einen attraktiven Analyseansatz dar. Immer nämlich besteht das Erfordernis, über einen längeren Zeitraum hinweg die öffentliche Meinung zu beobachten. Die Struktur des befragten Personenkreises muss dabei allerdings im Zeitverlauf konstant bleiben; ansonsten wäre es nicht eindeutig möglich, beobachtete Änderungen im Antwortverhalten als tatsächliche Änderungen der öffentlichen Meinung zu interpretieren, da sie ebenso Folge einer veränderten Struktur des befragten Personenkreises sein könnten. Dies ist allerdings bei herkömmlichen Umfragen der Fall: Stichproben gehen en bloc ins Feld, einige Zielpersonen antworten früher, andere später. Unterscheiden sich dann aber früh von spät gegebenen Antworten, könnte dies auf veränderte Rah-

menbedingungen oder aber auf strukturell verschiedene Antwortmuster hindeuten. Nur im ersten Fall wäre die Diagnose einer veränderten öffentlichen Meinung real; im zweiten Fall wäre sie ein Artefakt. An dieser Stelle setzt die Grundidee von RCS-Erhebungen an: Sie besteht darin, die Durchführung einer auf einer Zufallsstichprobe basierenden Querschnittsbefragung in strikt kontrollierter Weise so über einen vorab definierten Zeitraum zu spreizen, dass nicht nur alle Befragten zusammen, *sondern auch die Befragten jedes einzelnen Tages* jeweils in sich eine Zufallsstichprobe aus der Grundgesamtheit konstituieren.

Anlässlich der Bundestagswahl 2005 wurde die erste Umfrage dieses Typs in Deutschland realisiert. Ziel dieses Projektes ist die Analyse der dynamischen Wirkungen des Wahlkampfes zur vorgezogenen Bundestagswahl 2005 auf die Orientierungen der Wähler und darüber vermittelt auf das Wahlergebnis (Schmitt-Beck 2005). Der vorliegende Beitrag schildert die in diesem Projekt gemachten Erfahrungen. Idealtypisch betrachtet von bestechender Eleganz und Erklärungskraft, konfrontiert das RCS-Design den Forscher in der Praxis mit erheblichen Umsetzungsproblemen, auf welche die Literatur bislang kaum eingegangen ist. Wir werden im Folgenden zunächst idealtypisch beschreiben, worin die besonderen Eigenschaften von RCS-Studien bestehen. Anschließend wird erläutert, wie dieses anspruchsvolle Design in unserer Studie praktisch realisiert wurde. Dabei wird auch diskutiert, welche Aspekte sich als besonders schwierig erwiesen haben. Der Beitrag schließt mit einigen exemplarischen Analysen, welche Rückschlüsse auf die Datenqualität ermöglichen und die besonderen Vorzüge von RCS-Daten verdeutlichen.

2 Wahlkampfforschung und der *Rolling Cross-Section Survey*

Die Analyse der Wirkungen von Wahlkämpfen ist ein Spezialisierungsgebiet der Politikwissenschaft, das lange ein Schattendasein fristete und erst in den letzten zehn Jahren einen deutlichen Aufschwung genommen hat (Holbrook 1996; Farrell & Schmitt-Beck 2002; Brady & Johnston 2006; für einen aktuellen Überblick siehe Schoen 2005: 521-541). Das hat mit der lange vorherrschenden Vorstellung zu tun, dass Wahlkämpfe keinen wesentlichen Beitrag zur Erklärung von Wahlergebnissen leisten könnten, aber auch mit den besonderen methodischen Herausforderungen dieses Forschungsfeldes (Brady u. a. 2006).

Wahlkämpfe können als Ketten von Ereignissen konzeptualisiert werden (Hillygus & Jackman 2003: 584; Stöss 1997: 104). Diese werden überwiegend von Parteien und Kandidaten initiiert, können aber auch andere Hintergründe haben, etwa Publikationsentscheidungen von Medien, Handlungen anderer politischer Akteure (z. B. Interessengruppen), Entwicklungen der internationalen Politik oder sogar zunächst unpolitische Vorkommnisse wie z. B. Naturkatastrophen. Alle diese Ereignisse können sich auf die Entwicklung der Unterstützung von Parteien und Kandidaten auswirken, sofern sie Informationen transpor-

tieren, die zwischen diesen in wertender Weise diskriminieren und für die breite Öffentlichkeit wahrnehmbar werden. Dies wird in erster Linie durch die Vermittlung der Massenmedien erreicht (Holbrook 1996; Shaw 2000). Aufgabe der Wahlkampfforschung muss es demnach sein, Veränderungen von Einstellungen, Wahrnehmungen und Verhaltensabsichten sichtbar zu machen und sie mit Ereignissen des Wahlkampfes als möglichen Ursachen in Verbindung zu bringen. So können die Hintergründe der Dynamik der öffentlichen Meinung im Verlauf von Wahlkämpfen erklärt werden (Holbrook 1996: 48). Studien zur Analyse von Wahlkampfwirkungen müssen also längsschnittlich angelegt werden. „Because campaigns are dynamic phenomena, good campaign studies must be dynamic, too. Time must enter the analysis either directly or indirectly (as proxy for campaign activities and events).“ (Bartels 2006: 2; zit. n. d. Manuskript)

Einer der bedeutendsten methodischen Fortschritte in diesem Bereich war die Entwicklung des Panel-Designs mehrfacher Wiederholungsbefragungen derselben Personen durch Paul Lazarsfeld und seine Mitarbeiter Mitte des vergangenen Jahrhunderts. Die bahnbrechende Erie County-Studie zum Präsidentschaftswahlkampf 1940, bei der eine Stichprobe der Einwohner eines Landkreises in Ohio siebenmal in monatlichen Abständen befragt wurde (Lazarsfeld u. a. 1944), stellt eine bis zum heutigen Tag eindrucksvolle Pionierleistung dar. Panel-Studien sind seither die gebräuchlichste Methode, um dynamische Prozesse kurzfristiger Meinungsentwicklungen einzufangen. Panel-Designs haben den großen Vorteil, Vorgänge individuellen Wandels nachvollziehbar zu machen. Allerdings weisen sie auch einige gravierende Nachteile auf. Dazu zählen Konditionierungseffekte, aber auch die unvermeidliche Panel-Mortalität. Für Wahlkampfforscher ist insbesondere die zeitliche „Grobkörnigkeit“ solcher Erhebungen ein Problem. Aus logistischen Gründen ist es schwierig, zahlreiche Panel-Wellen in sehr dichter Folge zu realisieren.¹ Je mehr Zeit aber zwischen den einzelnen Panel-Wellen verstreicht, desto schwieriger ist es, Panel-Effekte eindeutig einem bestimmten Ereignis zuzuschreiben, weil die Zahl der potenziell für Effekte verantwortlichen Ereignisse mit der Zeit steigt. Außerdem weist das Instrument hinsichtlich möglicher Untersuchungsgegenstände eine gravierende Einschränkung auf. Der typische Untersuchungsplan, um Wirkungen eines Ereignisses mittels Panel-Daten zu analysieren, ist das Pre-Post-Design, d. h. die Durchführung von Befragungen vor und nach einem mutmaßlich einflussreichen Ereignis. Das ist aber nur möglich, wenn vor Studienbeginn schon entschieden wird, welches Ereignis untersucht werden soll. Nicht antizipierbare Vorkommnisse sind mit diesem Instrument ebenso wenig analysierbar

1 Mit der Verbreitung neuer Medien und dem Aufbau von Pools befragungswilliger Personen bieten sich allerdings auch in diesem Bereich neue Möglichkeiten, die sich etwa das OmniNet von FORSA zunutze macht. Im Vorfeld der Bundestagswahl 2002 wurde damit ein elf-welliges Panel realisiert (vgl. Güllner et al. 2005).

wie Ereignisse, die zwar im Prinzip antizipierbar sein mögen, für die aber mangels einer geeigneten Theorie keine Erwartungen politischer Effekte existieren. Für die in einer solchen Situation erforderlichen explorativen Analysen sind reine Panel-Designs ungeeignet (Johnston & Brady 2002; Romer u. a. 2004; Brady & Johnston 2006; Brady u. a. 2006; Bartels 1999, 2006).

Mit dem *Rolling Cross-Section-Design* wurde in den letzten Jahren eine neue Methode zur Analyse zeitlichen Wandels von Einstellungen und anderen Orientierungen entwickelt und konzeptionell perfektioniert, welche diese Probleme nicht aufweist (Johnston 2001; Johnston & Brady 2002; Brady & Johnston 2006; Brady u. a. 2006; Romer u. a. 2004). Amerikanische Kampagnenberater setzen Verfahren mit ähnlicher Logik schon länger (unter der Bezeichnung „tracking polls“) ein, um ihren Klienten möglichst detaillierte Daten über die Resonanz ihrer Wahlkampfmaßnahmen zu verschaffen (Blakeman 1995; Asher 1998: 112-3). Dies setzt eine *kontinuierliche* repräsentative Beobachtung der Zielgruppe voraus, was herkömmliche Umfragen in der Regel nicht leisten. Dort nämlich werden bestimmte Zielpersonen eher zu Beginn der Feldzeit, andere dagegen erst zu einem späteren Zeitpunkt erreicht und interviewt. Idealerweise aber müssen die Befragten eines jeden Tages jeweils in sich eine Zufallsstichprobe der Grundgesamtheit darstellen. Genau dies leisten RCS-Studien. Um die Beobachtung der öffentlichen Meinung im Zeitverlauf noch weiter zu standardisieren, wird darüber hinaus versucht, jeden Tag ungefähr gleich viele Interviews zu realisieren. Die für die inferenzstatistische Absicherung von Befunden zwingend notwendige Zufallsauswahl von Befragten kommt bei RCS-Studien auf doppelte Weise zum Tragen: Erstens im Hinblick auf die Chance eines Mitglieds der Grundgesamtheit, überhaupt in die Studie einbezogen zu werden, zweitens im Hinblick auf den Zeitpunkt innerhalb des Erhebungszeitraumes, an dem eine ausgewählte Befragungsperson interviewt wird. Zentrale Parameter, um diese distinktiven Merkmale von RCS-Studien zu gewährleisten, sind die Stichprobenziehung und die Feldadministration:

- Die Stichprobenziehung (präziser: die Ziehung der Haushaltsstichprobe) erfolgt bei einer RCS-Studie in zwei Stufen. Zunächst wird eine normale Zufallsstichprobe gezogen. Diese wird dann jedoch nicht wie üblich sofort komplett oder in unsystematischer – oftmals rein feldlogistischen Erwägungen folgender – Weise über einen gewissen Zeitraum freigegeben. Vielmehr wird sie in einem zweiten Schritt *zufällig* in Teilstichproben (so genannte *Replikate*) unterteilt, deren Anzahl der Zahl der geplanten Erhebungstage entspricht. Jedes Replikat stellt ebenfalls eine zufällige Stichprobe aus der Grundgesamtheit dar. Diese werden dann nach einem festgelegten „Fahrplan“ jeweils an einem bestimmten, zufällig ausgewählten Tag für die Befragung freigegeben.

- Nach der Freigabe wird für jedes Replikat eine Bearbeitungsroutine in Gang gesetzt, die einem strengen Protokoll folgt. Diese hat das Ziel, durch intensive Feldarbeit jede Teilstichprobe möglichst gut auszuschöpfen. Dies ist notwendig, weil in der Feldpraxis nicht alle Mitglieder eines Replikats umgehend am Tag der Freigabe interviewt werden können. Das wird nur für leicht erreichbare (und befragungsbereite) Personen der Fall sein, nicht jedoch für schwerer erreichbare Personen. Mit diesen kommen Interviews typischerweise erst nach mehreren vergeblichen Kontaktversuchen zustande. Essentiell für das RCS-Design ist nun, dass dieses Protokoll für alle Replikate in strikt einheitlicher Weise abgearbeitet wird. Um die Gleichheit aller Replikate im Hinblick auf ihre Repräsentativität für die Grundgesamtheit und die dafür notwendige Eigenschaft der zufälligen Auswahl zu gewährleisten, muss jedes von ihnen nach exakt denselben Kontaktierungsregeln bearbeitet werden. Weder der Start- oder Wochentag noch sonstige Gesichtspunkte dürfen dabei eine Rolle spielen. Diese Regeln beinhalten einen genau festgelegten, für alle Replikate gleich langen, mehrtägigen Zeitraum, für den gemeinsame Adressen bzw. Telefonnummern eines Replikats aktiv bleiben und Kontaktversuche unternommen werden. Auch für letztere gilt im Hinblick auf Anzahl und Zeitpunkte eine präzise Routine. Würden diese Abläufe geändert, würden nicht mehr alle Mitglieder der Grundgesamtheit mit derselben Wahrscheinlichkeit in die Stichprobe inkludiert und einem bestimmten Replikat alloziert. Damit könnte die für das RCS-Design zentrale Annahme der völligen strukturellen Gleichartigkeit aller Replikate nicht mehr aufrechterhalten werden.²

Die dichte zeitliche Taktung von RCS-Studien auf Tagesbasis und die Notwendigkeit strikter Kontrolle über den Befragungszeitpunkt präjudiziert den Erhebungsmodus: tägliche RCS-Erhebungen können nur telefonisch durchgeführt werden.³ Bei genauer Umsetzung der beschriebenen methodischen Anforderungen erhält man eine Umfrage, die alle Eigenschaften einer normalen Querschnittsstudie auf der Grundlage einer Zufallsstichprobe besitzt und auch als solche analysiert werden kann. Zugleich aber kann sie unter zeitli-

2 Dies impliziert, dass die zeitliche „Schrittgröße“ der Erhebungen während der gesamten Feldzeit nicht variiert werden darf. RCS-Studien müssen nicht zwingend auf täglicher Basis erfolgen, denkbar sind beispielsweise auch Erhebungen im Wochenrhythmus. Allerdings ist zu bedenken, dass größere Abstände zwischen Replikaten auf Kosten der Genauigkeit bei der Analyse der Effekte von Ereignissen gehen, weil dann eine präzise Ursachenattribution zu ganz bestimmten Ereignissen erschwert wird. Völlig unvereinbar mit der Logik des RCS-Design sind aber ungleichmäßige oder auch systematisch variiende Abstände zwischen den Erhebungszeitpunkten. Ebenso ausgeschlossen ist die – bei Standardbefragungen übliche – Beendigung der Befragung nach Erreichen einer bestimmten Fallzahl.

3 Es gibt zwar persönlich-mündlich durchgeführte RCS-Studien, beispielsweise die Niederländische Wahlstudie 1998, doch arbeitete diese mit wöchentlicher Taktung.

chen Aspekten in beliebiger und völlig flexibler Weise in ebenfalls zufällige Unterstichproben für beliebige Zeitpunkte oder -perioden zerlegt werden. Durch diese Eigenschaft von RCS-Daten werden höchst differenzierte Analysen des dynamischen Geschehens im Verlauf von Wahlkämpfen möglich. Die Replikate, die bei der Erhebung von RCS-Daten zentrale Bedeutung haben, treten bei diesen Analysen nicht mehr in Erscheinung. Die Auswertungen orientieren sich vielmehr an den Kalendertagen, an denen die Interviews realisiert werden. Auf ihrer Grundlage können „synthetische Querschnittsstichproben“ gebildet werden. Diese beinhalten alle Interviews, die innerhalb eines vom Forscher definierten Zeitraums durchgeführt wurden.

Abbildung 1 Zeitliche Struktur eines Rolling Cross-Section Survey

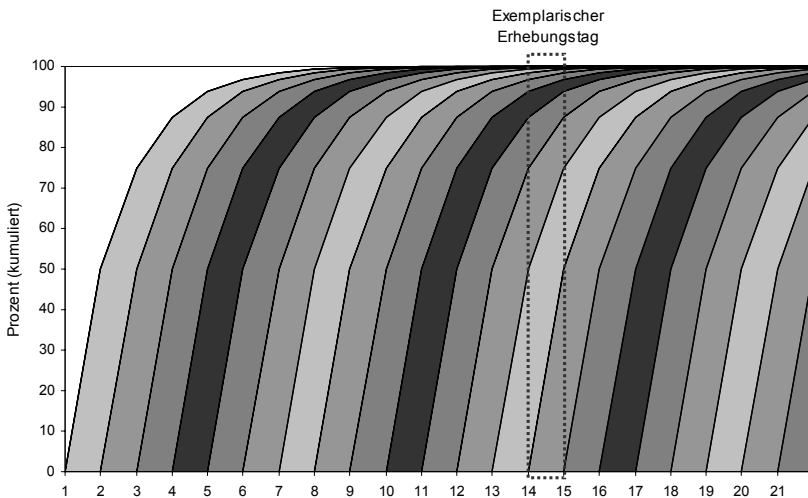


Abbildung 1 verdeutlicht die idealtypische Grundstruktur der durch RCS-Erhebungen gewonnenen Daten.⁴ Die Logik des RCS-Designs gewährleistet, dass die für jeden Erhebungstag zur Verfügung stehenden Interviews jeweils identische Mischungen von Befragten aus mehreren Replikaten und damit von leicht, aber auch schwer erreichbaren Personen beinhalten. Mit Blick auf die Stichprobe insgesamt ist der Befragungstag innerhalb

4 Bei der Darstellung handelt es sich um eine simulierte Datenstruktur unter der Maßgabe, dass jeden Tag das zu diesem Zeitpunkt in jedem Replikate noch vorhandene Potenzial noch nicht realisierter Interviews zu jeweils 50 Prozent ausgeschöpft wird.

der Feldzeit damit nicht mehr systematisch, sondern rein zufällig bedingt und besitzt somit den Charakter eines Zufallsereignisses. Die an den verschiedenen Erhebungstagen realisierten Interviews unterscheiden sich durch nichts außer dem Datum ihrer Durchführung. Noch präziser formuliert unterscheiden sie sich nur hinsichtlich des situativen Kontextes im Hinblick auf das Wahlkampfgeschehen, das diesem Datum entspricht. Alle Änderungen von Wahrnehmungen, Einstellungen, Präferenzen oder Verhaltensorientierungen, die im Zeitverlauf identifiziert werden, können daher nur zwei Quellen haben: den zufälligen Stichprobenfehler und realen Wandel unter dem Eindruck des Kampagnengeschehens. Diese können durch geeignete statistische Verfahren separiert werden (Johnston & Brady 2002; Brady & Johnston 2006; Romer u. a. 2004).

Auf diesem Wege gewonnene Daten können in äußerst flexibler und vielfältiger Weise analysiert werden. Man kann die Daten auf der Basis einzelner Erhebungstage analysieren und damit punktgenaue Verbindungen zwischen Wahlkampfereignissen und zugehörigen Änderungen der öffentlichen Meinung herstellen.⁵ Auch die Dauerhaftigkeit solcher Veränderungen ist leicht zu beobachten.

Neben der tageweisen Betrachtung können die Daten ebenso nach Formalkriterien zusammengefasst werden, beispielsweise wochenweise, um zu höheren Fallzahlen und damit höherer Teststärke zu gelangen. Da die Daten gleichmäßig über den gesamten Untersuchungszeitraum verteilt sind, können Schnittpunkte aber auch in beliebiger Weise nach inhaltlichen Kriterien gesetzt werden, etwa in Form eines Vergleiches der Interviews unmittelbar vor und unmittelbar nach einem bestimmten Ereignis im Wahlkampf. Ein besonderer Vorzug ist dabei, dass dies keineswegs nur für antizipierbare Ereignisse wie z. B. Parteitage oder TV-Duelle gilt, sondern für jedes beliebige Ereignis, einschließlich solcher, die im Vorfeld überhaupt nicht vorhergesehen werden konnten. Letztlich ist jede beliebige Kumulation von Daten aufeinanderfolgender Erhebungstage möglich. „From the sampling perspective, all that adding or subtracting a day does is reduce or expand standard error. [...] So the design combines flexibility in periodization with power in combination of days.” (Brady & Johnston 2006: 12; zit. n. d. Manuskript)⁶

5 Dass sich über das Datum als Schlüssel externe Daten problemlos zuspielen lassen, beispielsweise aus Inhaltsanalysen der tagesaktuellen Medienberichterstattung, erleichtert auch die inhaltliche Interpretation beobachteter Änderungen (Gidengil & Dobrzynska 2003; Dobrzynska u. a. 2003; Fournier u. a. 2004).

6 Gelegentlich findet man in der Literatur auch Analysen normaler Querschnittsdaten, welche das Erhebungsdatum als unabhängige Variable einsetzen; allerdings ist das methodisch fragwürdig, weil bei solchen Studien der Befragungszeitpunkt mit der Erreichbarkeit der Befragten konfundiert ist, die ihrerseits häufig mit inhaltlich relevanten Merkmalen korreliert.

Nicht übersehen werden kann allerdings, dass RCS-Umfragen im Vergleich zu Panel-Befragungen auch einen deutlichen Nachteil aufweisen: Sie erlauben keine Aussagen über individuellen Wandel, sondern nur Aggregataussagen über Prozesse des Wandels in der gesamten Wählerschaft oder in Wählergruppen. Von daher bietet es sich an, RCS-Studien durch eine Panel-Komponente anzureichern, welche genau dies ermöglicht. Panel- und RCS-Studien können als komplementär gelten und ergänzen einander. Die Kombination einer RCS-Erhebung als Vorwahlbefragung mit einer als Nachwahlbefragung konzipierten zweiten Panel-Welle hat sich in den letzten Jahren als ideales Verfahren erwiesen, um die Kampagnendynamik bei Wahlkämpfen zu analysieren (Johnston 2001; Brady & Johnston 2006; Bartels 2006). Auch die hier beschriebene RCS-Erhebung wurde durch eine als Nachwahlbefragung realisierte zweite Panel-Welle angereichert. Die Variablen der Nachwahl-Welle können als Messungen des kumulativen Einflusses des gesamten Wahlkampfes interpretiert und als solche mit unterschiedlich weit in der Wahlkampffperiode zurückliegenden Vorwahl-Messungen derselben Variablen verglichen werden. Überdies lassen sich verschiedene für die Wahlkampfdynamik wesentliche Daten erst nach der Wahl sinnvoll erheben, nicht zuletzt natürlich die tatsächliche Wahlentscheidung selbst.

3 Die RCS-Studie zur Bundestagswahl 2005

3.1 Grundkonzept

Die erste RCS-Studie wurde als Pilotprojekt im Rahmen der American National Election Study (ANES) 1984 realisiert (Bartels 1988; Allsop & Weisberg 1988; Whiteley 1988). Allerdings erfolgte die Datenerhebung hier noch in wöchentlichem Turnus. Der eigentliche Durchbruch des RCS-Designs setzte die Umstellung auf tägliche Interviews voraus, welche erstmals im Rahmen der Canadian Election Study (CES) 1988 vorgenommen wurde. Seitdem ist das (mit einer Nachwahl-Panel-Welle kombinierte) RCS-Design bei der CES Standard (Johnston u. a. 1992, 1996; Nevitte u. a. 2000; Blais u. a. 2002). 1996 übernahm die nationale Wahlstudie Neuseelands (NZES) dasselbe Modell (Vowles u. a. 1998, 2002, 2004), seit 2001 beinhaltet auch die britische nationale Wahlstudie (BES) eine RCS-Komponente (Sanders et al. 2006; Clarke u. a. 2004, 2006). Die niederländische nationale Wahlstudie 1998 enthielt ebenfalls eine RCS-Komponente, jedoch mit wöchentlichem Erhebungsturnus (Aarts u. a. 1999). Eine RCS-Befragung konstituiert auch den Kern des größten jemals realisierten Forschungsprogramms der Wahlsoziologie: Der anlässlich der US-Präsidentenwahlen 2000 und 2004 durchgeführte, wahrhaft monumentale National Annenberg Election Survey (NAES) umfasst bis zu 300 tägliche Interviews über einen Zeitraum von 14 Monaten hinweg (Romer u. a. 2004; Johnston u. a. 2004; Clymer 2006). Schließlich wurden auch für die Bundestagswahlen 1994 und 2002

Analysen vorgelegt, die auf einem RCS-Umfragen ähnlichen Design basieren, den seit 1991 von FORSA durchgeführten kommerziellen OmniTel-Tagesbefragungen (Stöss 1997; Erhardt 1998; Dörmer 1998; Güllner u. a. 2005). Diese haben allerdings den Nachteil eines uneinheitlichen Erhebungsrhythmus: Nur an Werktagen werden Befragungen durchgeführt (Güllner 2000), was den idealtypischen Anforderungen eines RCS-Designs widerspricht. Ob es noch andere Abweichungen gibt, kann nicht beurteilt werden, da die genauen Details der Feldarbeit nach unserer Kenntnis nicht wissenschaftsöffentlich dokumentiert sind.

Im Vorfeld der vorgezogenen Bundestagswahl 2005 wurde die erste RCS-Studie in Reinform als Wahlstudie in Deutschland realisiert. Ziel des Projektes ist die Analyse der dynamischen Wirkungen des Wahlkampfes zu dieser ungewöhnlichen Wahl auf die Orientierungen der Wähler und darüber vermittelt auf das Wahlergebnis (Schmitt-Beck 2005). Im Folgenden beschreiben wir, wie das idealtypisch äußerst elegante, praktisch aber ungewöhnlich anspruchsvolle RCS-Design bei dieser Studie umgesetzt wurde und welche Erfahrungen wir dabei gemacht haben.⁷

Die RCS-Studie zur Bundestagswahl 2005 wurde in der Zeit vom 8. August bis zum 17. September 2005 durchgeführt; sie umfasste also einen Zeitraum von sechs Wochen und endete am letzten Tag vor der Bundestagswahl. Sie wurde als CATI-Erhebung realisiert und umfasst 3.583 zufällig ausgewählte Befragte aus der Grundgesamtheit der deutschsprachigen, in Privathaushalten mit mindestens einem Festnetzanschluss lebenden Bevölkerung der Bundesrepublik Deutschland, die zur Bundestagswahl 2005 wahlberechtigt war. Die Feldarbeit oblag dem Institut *Ipsos GmbH* (Möln). Die Stichprobenziehung erfolgte als ADM-Telefonstichprobe gemäß dem Gabler/Häder-Modell (Gabler & Häder 1997, Häder & Gabler 1998). Da eine Nachziehung von Telefonnummern designbedingt ausgeschlossen werden musste, war es notwendig, eine sehr umfangreiche Bruttostichprobe generierter Telefonnummern zu ziehen. Diese wurde dann in zufällige Teilstichproben zerlegt. Um ausreichend Reservetelefonnummern verfügbar zu haben, wurden insgesamt 60 Replikate zu je 600 Telefonnummern gezogen. Damit stand eine Bruttostichprobe von insgesamt 36.000 generierten Telefonnummern zur Verfügung, von denen letztlich 25.000 genutzt wurden.⁸ Angestrebt war zunächst, täglich 80 Interviews zu realisieren; nach 11 Tagen wurde diese Zielmenge auf 90 Interviews pro Tag erhöht.

7 Die hier beschriebene Studie wurde in engem Kontakt mit den Primärforschern von CES, NAES bzw. BES Richard Johnston, Elisabeth Gidengil, David Northrup und Harold Clarke geplant, denen wir für wertvolle Informationen und Anregungen zu großem Dank verpflichtet sind.

8 Diese sehr umfangreiche Bruttostichprobe wurde aus zwei Gründen für notwendig erachtet: Es war nicht bekannt, welches Brutto nötig war, um die angestrebten Interviewzahlen pro Tag zu erreichen, außerdem war zum Zeitpunkt der Planung der Studie noch nicht klar, ob die Bundes-

In unseren Vorüberlegungen hatten wir antizipiert, dass die Ausschöpfungsquoten nicht an allen Wochentagen gleich sein würden und möglicherweise auch in Abhängigkeit von nicht antizipierbaren Faktoren schwanken könnten. Es war daher notwendig, bei der Stichprobenziehung Voraussetzungen zu schaffen, die es erlauben würden, die Teilstichproben kontinuierlich mit dem Ziel möglichst gleichmäßiger täglicher Fallzahlen genauer auszusteuern. Dabei durfte allerdings die Logik des RCS-Designs nicht aufgegeben werden. Die dafür entwickelte Lösung sah vor, die Teilstichproben – wiederum zufällig – noch weiter in „Scheiben“ zu je 100 Telefonnummern zu zerlegen. Auch diese Scheiben waren aufgrund ihrer zufälligen Ziehung beliebig kumulierbar; es handelte sich gleichsam um „Replikate 2. Ordnung“. Dadurch wurde die Möglichkeit geschaffen, die Teilstichproben in einer dem RCS-Design angemessenen Weise ad hoc durch Hinzuspielen oder Weglassen einzelner Scheiben zu vergrößern oder zu verkleinern, um so die täglichen Fallzahlen realisierter Interviews möglichst konstant zu halten. Die Festlegung der Zahl der an jedem Erhebungstag neu einzuspielenden Scheiben erfolgte jeweils zeitnah in Reaktion auf die aktuelle Feldentwicklung und die bis dahin gesammelten Erfahrungen mit dem Verlauf der Erhebung. Alle Teilstichproben und Scheiben erhielten individuelle Kennungen; für jeden einzelnen Fall im Datensatz können daher Interviewzeitpunkt, Teilstichprobe bzw. Replikat sowie Scheibe innerhalb dieser Teilstichprobe eindeutig bestimmt werden.

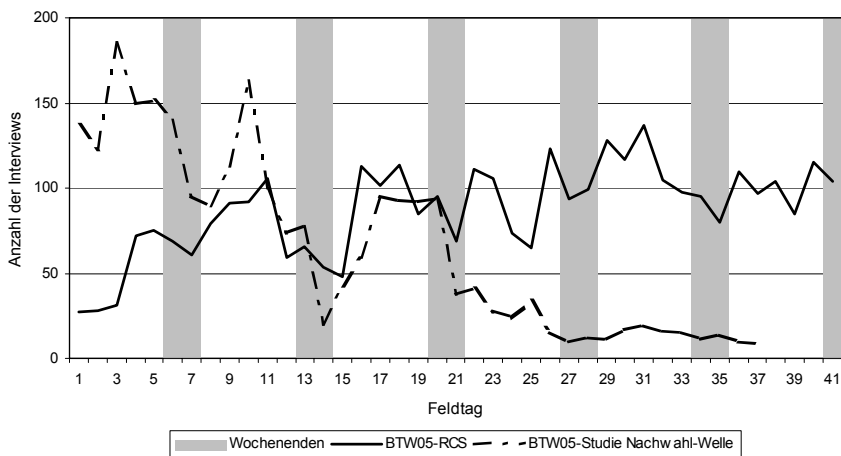
3.2 Feldverlauf

RCS-Studien können in der Erhebungsphase keinesfalls „sich selbst überlassen“ werden. Die Feldentwicklung bedarf permanenter Beobachtung und gegebenenfalls zeitnaher Adjustierung, sonst besteht ein erhebliches Risiko, das Ziel eines gleichmäßigen Feldverlaufs auf dem angestrebten Niveau täglicher Fallzahlen zu verfehlen. Faktisch wurde im Feldverlauf der Bundestagswahlstudie fast immer mit allen 600 Nummern, d. h. 6 Scheiben gearbeitet. An einem Tag wurde aufgrund sehr guter Feldentwicklung versuchsweise auf eine Scheibe verzichtet. Diese Maßnahme führte zu einer deutlichen Verringerung der Fallzahl und wurde daher am Folgetag wieder zurückgenommen. An einigen anderen Tagen wurden zusätzliche (7.) Scheiben eingesetzt, um absehbar geringere Fallzahlen zu kompensieren. Zumeist betraf dies Wochenenden.

tagswahl tatsächlich am 18. September 2005 stattfinden würde, weil noch die Entscheidungen des Bundespräsidenten und des Bundesverfassungsgerichts ausstanden. Die Zahl der Teilstichproben wurde in Anbetracht der nicht ausschließbaren Möglichkeit eines späteren Wahltermins bewusst sehr hoch angesetzt.

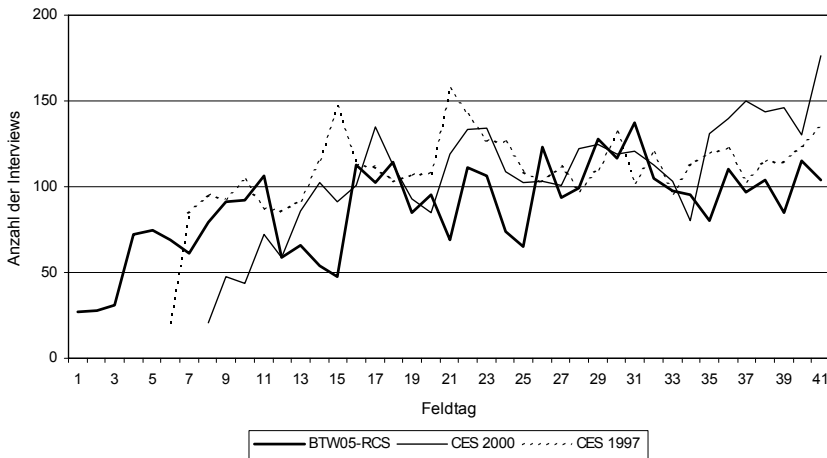
Wie sich der Feldverlauf der Bundestagswahlstudie entwickelte, ist Abbildung 2 zu entnehmen. Sie gibt die Zahl der Interviews wieder, die an jedem Feldtag realisiert wurden. Zum Vergleich ist auch der Feldverlauf der Nachwahl-Welle unseres Projektes eingezeichnet, der exemplarisch für die Entwicklung einer typischen Querschnittsbefragung stehen kann. Hier war die Maxime, möglichst viele Interviews in möglichst kurzer Zeit zu realisieren. Die Feldentwicklung der Nachwahlstudie verlief zwar nicht völlig monoton, insbesondere sind Einbrüche an den Wochenenden, und da wiederum vor allem sonntags, deutlich zu erkennen. Aber der Gesamttrend entwickelte sich klar von sehr hohen Zahlen realisierter Interviews in den ersten Tagen der Feldperiode (die sofort am Montag nach der Bundestagswahl einsetzte) zu geringeren täglichen Ausbeuten und schließlich einer Endphase mit schleppendem Feldverlauf auf niedrigem Niveau. Im Vergleich dazu stellt sich, als Folge der systematischen Zeitsteuerung, der Feldverlauf der RCS-Welle sehr viel gleichmäßiger dar.

Abbildung 2 Zahl der täglichen Interviews: RCS-Studie zur Bundestagswahl 2005 im Vergleich zur Nachwahl-Welle



Quelle: RCS-Studie „Kampagnendynamik 2005“

Abbildung 3 **Zahl der täglichen Interviews: RCS-Studie zur Bundestagswahl 2005 im Vergleich zur CES 1997 und 2000**



Quelle: RCS-Studie „Kampagnedynamik 2005“, Canadian Election Studies 1997 und 2000 (vgl. Northrup 1998: 7, 2003: 11)

Gemessen am Ziel einer möglichst gleichmäßigen Zahl täglicher Interviews mag der Feldverlauf indessen immer noch recht unregelmäßig erscheinen. So wird erkennbar, dass die ersten Tage deutlich hinter den Zielzahlen zurückblieben. Das ist zum Teil eine logische Konsequenz der Stichprobenallokation, denn die angestrebten täglichen Fallzahlen können erst dann erreicht werden, wenn eine größere Zahl von Replikaten gleichzeitig aktiv ist, was unmittelbar nach Feldbeginn natürlich noch nicht der Fall sein konnte. Ein gewisses „Aufschaukeln“ am Anfang ergibt sich zwangsläufig aus der Logik des RCS-Designs und zeigt sich analog auch bei vergleichbaren anderen Studien. Das belegt Abbildung 3 am Beispiel der kanadischen Wahlstudien 1997 und 2000. Eine wichtige Implikation hiervon ist im Übrigen, dass die Interviews der ersten Erhebungstage noch keine perfekten Zufallsstichproben darstellen und daher nicht in dynamische Analysen einbezogen werden können (Johnston & Brady 2002: 285). Das ist auch in der idealtypischen Darstellung von Abbildung 1 erkennbar. Im Fall der deutschen Studie kamen aber in den ersten Feldtagen Probleme mit dem Adressenverwaltungssystem hinzu, auf die weiter unten zurückzukommen ist.

Festzustellen ist aber auch ein insgesamt nicht völlig ebener Feldverlauf mit größeren Tagesschwankungen. Das betrifft zunächst einmal die Wochenenden, und da wiederum

vor allem die Sonntage, an denen – wie bei der Nachwählerhebung – stets weniger Interviews realisiert werden konnten als an den Werktagen zuvor und danach. Wie beschrieben, wurde dieser Tendenz teilweise durch den Einsatz zusätzlicher Scheiben an den Wochenenden entgegengewirkt. Das hat verhindert, dass diese Einbrüche noch größer wurden. Aber solche Interventionen sind nicht unbegrenzt möglich, weil sonst designbedingte „Aufschaukelungsprozesse“ an den Folgetagen ausgelöst würden, an denen diese Scheiben ja ebenfalls aktiv bleiben. Daher konnten die „Wochenendtäler“ aus systematischen Gründen nicht vollständig nivelliert werden. Durchaus erhebliche, aber nicht systematisch deutbare Unterschiede zeigen sich überdies auch zwischen den Werktagen.⁹ Das ist allerdings auch bei den beiden kanadischen Vergleichsstudien nicht anders.¹⁰ Die Fallzahlen täglich realisierter Interviews schwanken bei der Bundestagswahlstudie zwischen 27 am ersten Feldtag und 137 am 31. Feldtag bei einem Gesamtdurchschnitt von 87.4 Interviews pro Tag und einer Standardabweichung von 26.7. Erwähnung verdient auch ein Phänomen, das sich in allen kanadischen RCS-Studien zeigte (siehe auch Johnston & Brady 2002: 285), nicht jedoch in der deutschen: eine – vermutlich seitens der Befragten motivationsbedingte – Zunahme der Ausschöpfung der Replikate in den letzten Tagen vor der Wahl, die sich in einem deutlichen Anstieg der realisierten Fallzahlen niederschlug. Die angestrebten täglichen Fallzahlen erweisen sich somit in der Praxis als Richtgröße, während die Zahl faktisch realisierter Interviews offenkundig unvermeidbaren Fluktuationen unterworfen ist – ein Phänomen, das alle RCS-Studien betrifft. So präzise, dass ein völlig ebener Feldverlauf erzielt wird, lässt sich der Interviewprozess bei RCS-Studien nicht steuern, weil nur Instrumente zur Feinsteuerung eingesetzt werden können, welche die Erfordernisse des RCS-Designs nicht verletzen.

3.3 Bearbeitung der Replikate

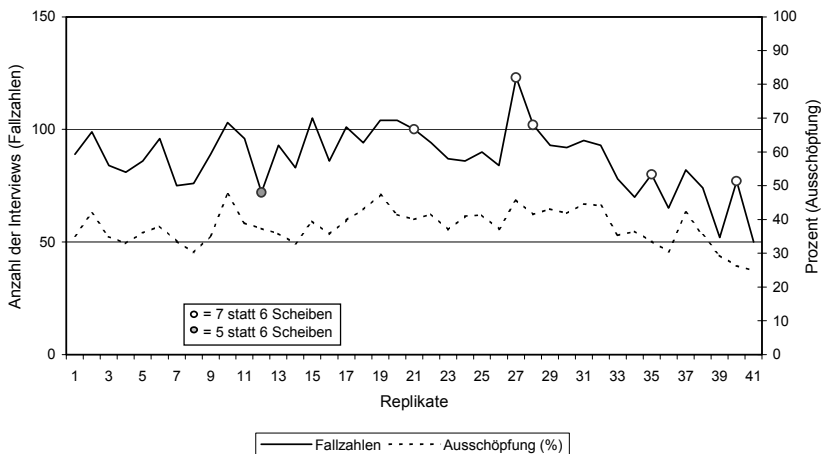
Neben der tageweisen Betrachtung ist auch eine Betrachtung der Replikate, die an jeweils einem Tag ins Feld gegeben und an diesem und den folgenden Tagen abgearbeitet wurden, möglich. Abbildung 4 zeigt die Zahl der Interviews, die aus jedem der 41 täglichen Replikate hervorgegangen sind. Zwischen den Teilstichproben sind deutlich weniger Schwan-

9 Ein Einbruch im Verlauf der Felddurchführung ist klar auf technische Probleme zurückführbar: Am 15. Erhebungstag konnten ab ca. 15:00 keine Interviews mehr durchgeführt werden, weil bei Bauarbeiten in der Umgebung des Telefonstudios ein Stromkabel durchtrennt und damit das gesamte Telefonstudio „lahm gelegt“ worden war – ein anschauliches Beispiel für die Unwägbarkeiten, welche die Umsetzung eines methodisch rigorosen Designs wie des hier beschriebenen beeinträchtigen können.

10 Für andere RCS-Studien sind leider keine numerischen Detailinformationen über täglich realisierte Fallzahlen dokumentiert; siehe jedoch auch die ganz ähnlichen Grafiken zur CES 1993 bei Johnston (2001: 155) und zur NAES 2000 bei Romer (2004: 58).

kungen zu verzeichnen als zwischen den Erhebungstagen (s. o.). Auffällig ist vor allem ein einsamer Gipfel für ein Replikat, aus dem 123 Interviews realisiert werden konnten. Zumindest teilweise ist das dadurch zu erklären, dass es sich hierbei um eine der Teilstichproben mit erhöhter Scheibenzahl, d. h. größerem Bruttoumfang, handelte. Auch einige der anderen „Gipfel“ sind auf erhöhte Scheibenzahlen zurückzuführen, eines der „Täler“ wurde korrespondierend durch Reduktion der Scheibenzahl bewirkt. Das standardisierte Bild der Ausschöpfungsquoten der Teilstichproben ist demgegenüber sehr viel gleichmäßiger. Auffällig ist die stetig schrumpfende Zahl der Interviews, die aus den letzten Teilstichproben realisiert wurden – ein Muster, das auch in anderen RCS-Studien festgestellt wurde. Dieses Phänomen ist designbedingt: Die letzten vor der Wahl ins Feld gegebenen Replikate konnten nicht für dieselbe Zeitdauer offen bleiben wie diejenigen, die so rechtzeitig ins Feld gegangen waren, dass sie noch vor der Wahl vollständig „abgearbeitet“ werden konnten. Da die Erhebung am Tag vor der Wahl eingestellt wurde „trunkierte“ der Wahltag diese Replikate und setzte ihrer Bearbeitung gleichsam ein natürliches Ende. Das ist idealtypisch auch aus Abbildung 1 ersichtlich. Für die späteren Analyse-möglichkeiten ist das jedoch unerheblich, denn die Stichprobenqualität der vorliegenden *Erhebungstage* – die letztlich entscheidend sind – wird dadurch nicht beeinträchtigt.

Abbildung 4 Realisierte Fallzahlen und Ausschöpfung aus Replikaten



Quelle: RCS-Studie „Kampagnendynamik 2005“

Die Daten der einzelnen Erhebungstage setzen sich jeweils aus Interviews zusammen, die sehr rasch nach Replikatöffnung durchgeführt wurden, und anderen, bei denen es länger, manchmal sogar sehr lange dauerte, bis sie realisiert werden konnten. Eine rigorose Feldbearbeitung mit dem Ziel, auch schwer erreichbare Befragte in die Studie einzubeziehen, ist kennzeichnend für RCS-Studien. Auch bei der deutschen Bundestagswahlstudie wurde auf eine außerordentlich intensive Feldarbeit Wert gelegt, um eine möglichst gute Ausschöpfung der einzelnen Replikate zu erreichen. Dazu gehörten beispielsweise gezielte Konversionsversuche „weicher“ Verweigerer¹¹: Sie wurden zwei Tage nach ihrer ursprünglichen Ablehnung noch einmal kontaktiert. Insgesamt 104 zusätzliche Interviews konnten so gewonnen werden, was einer Konversionsrate von 14.1 Prozent aus den ursprünglichen Verweigerungen entspricht. Die Interviewer konnten dabei, um ihre Argumentation zu unterstützen und glaubhaft zu machen, auf die Homepage von *Ipsos* verweisen, auf der eine Kurzbeschreibung des Projekts abgelegt war, sowie einen Link zu der ausführlicheren Projektdarstellung auf der Homepage der Primärforscher an der Universität Duisburg-Essen. Der Hinweis auf diese Informationsquelle konnte natürlich auch schon bei Befragten erfolgen, die in der ersten Kontaktpphase zögerlich waren.

Überdies wurden intensiv und über einen langen Zeitraum Kontaktversuche bei nicht erreichbaren, aber nicht als ungültig erkennbaren generierten Telefonnummern unternommen. Jedes Replikat blieb 21 Tage lang aktiv (um das Risiko zu minimieren, dass Kontaktversuche vollständig in Phasen urlaubsbedingter Abwesenheit möglicher Befragungspersonen fielen). In den ersten beiden Tagen sollte jede nicht sofort stichprobenneutral ausgeschiedene Telefonnummer, bei der noch kein Interview realisiert werden konnte, 8-mal angerufen werden, an den Folgetagen jeweils 2-mal. Im gesamten Feldverlauf nicht erreichbare Nummern waren somit insgesamt 54-mal anzurufen. Bezüglich der Uhrzeiten der Kontaktversuche war dabei ein komplexes Rotationsschema zu beachten. Telefoniert wurde jeden Tag von 10 bis 21 Uhr, sonntags von 14 bis 20 Uhr. Damit ging das Design der Bundestagswahlstudie noch einmal deutlich über die diesbezüglich auch schon sehr „aggressiven“ (Romer u. a. 2004: 59) anderen RCS-Studien hinaus.¹² Allerdings traten bei der Umsetzung dieses sehr aufwendigen Kontaktschemas Probleme auf. Diese hatten mit der Priorisierung der Kontaktversuche zu tun, die in den Algorithmus zur

11 Das sind Verweigerer, die z. B. unter der Begründung, keine Zeit oder kein Interesse zu haben, ein Interview ablehnten.

12 Das bei den CES angewandte Schema sah für jedes Replikat 2 bis 4 Kontaktversuche an den ersten vier Feldtagen und einen Kontaktversuch an den verbleibenden Feldtagen vor, und zwar für insgesamt 12 (CES 1988 und 1993) bzw. 10 Tage (CES 1997 und 2000). Bei der NAES 2000 erfolgten an den ersten vier Feldtagen je zwei Kontaktversuche, an den restlichen Tagen einer; jedes Replikat blieb 14 Tage aktiv (Romer u. a. 2004: 59).

Telefonnummernverwaltung einprogrammiert werden musste. Dabei erhielten Terminvereinbarungen die höchste Priorität¹³, gefolgt von besetzten Nummern (die nach 20 Minuten erneut zu kontaktieren waren), neu eingespielten Nummern aus dem aktuellen Replikat und schließlich den bereits eingesetzten Nummern aus den Replikaten der Vortage.

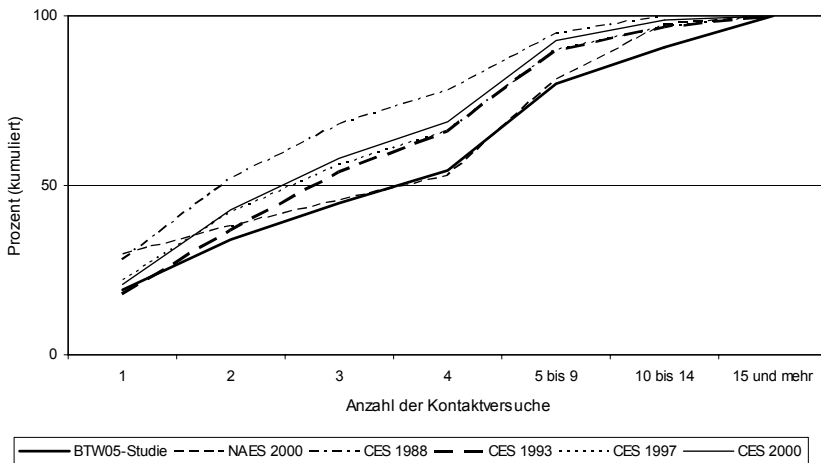
Vor dem Hintergrund der sehr großen Bruttostichprobe und der hohen Zahl von notwendigen Kontaktversuchen führte dieses Schema in der Praxis zu Klumpungen: Zu viele Nummern wurden zeitgleich ins System eingespielt, um sie alle tatsächlich kontaktieren zu können. In der Folge blieb die reale Zahl der Kontaktversuche unter dem angestrebten Soll. Als nach den ersten drei Feldtagen die realisierten Fallzahlen hinter den Erwartungen zurück blieben, wurde der Algorithmus geprüft und die Zeiten der Einspielung gleichmäßiger verteilt. In der Folge wurden dann die Sollfallzahlen rasch erreicht. Im Nachhinein zeigte jedoch eine aufwendige Analyse der Kontaktprotokolle, dass das angestrebte Kontaktschema auch danach nicht im vollen Umfang umgesetzt wurde. Dafür gab es während der Feldphase keine Indizien, weil es lediglich zu einer Verlangsamung des Feldverlaufes der einzelnen Teilstichproben führte, aber nicht zu einer sichtbaren Dämpfung der erreichten Fallzahlen. Wie sich bei der Nachanalyse zeigte, wurden die noch offenen Telefonnummern an den ersten beiden Öffnungstagen im Schnitt nur etwa halb so oft angerufen wie angestrebt (3.9- bzw. 3.8-mal statt 8-mal), an den Folgetagen (mit einem Soll von 2 Kontaktversuchen) lagen die durchschnittlichen Kontaktversuchsrate zwischen 2.3 und 1.4. Hieraus ist eine Lehre für künftige Studien dieses Typs abzuleiten: Der Entwicklung und Programmierung des Adressverwaltungsalgorithmus, der die Priorisierung und Abarbeitung der Kontaktversuche steuert, muss im Vorfeld größte Beachtung geschenkt werden. Aufgrund seiner zwangsläufigen Komplexität wäre im Grunde ein Pretest wünschenswert. Doch würde das nichts anderes bedeuten, als dass die Studie einige Tage unter Feldbedingungen laufen müsste, weil sich Kontaktierungsprobleme nur dann zeigen können. Diese äußerst aufwändige und im Grunde prohibitiv kostspielige Option wurde von uns im Vorfeld nicht ins Auge gefasst; unser Projekt beinhaltete nur einen normalen Pretest mit wenigen Interviews, welcher der Optimierung des Fragebogens diene.

Abbildung 5 zeigt im Vergleich mit allen CES-Studien und der NAES 2000, wie viele Kontaktversuche für die Realisierung der Interviews nötig waren. Dabei wird deutlich, dass sich eine intensive Feldarbeit lohnt. In keiner RCS-Studie wurde mehr als etwa ein Viertel der insgesamt am Ende realisierten Interviews schon beim ersten Kontaktversuch erzielt. Lediglich die CES 1988 kam bereits beim zweiten Kontaktversuch über die 50-

13 Streng genommen verstoßen Terminvereinbarungen mit Befragungspersonen gegen die strikte Logik des RCS-Designs, aber sie sind bei Erhebungen nur um den Preis einer deutlich verschlechterten Ausschöpfung verzichtbar.

Prozent-Marke. Im Vergleich zu den anderen Studien erweist sich der Feldverlauf bei der deutschen Studie als besonders langwierig. Die 50-Prozent-Marke wurde erst nach vier Kontaktversuchen erreicht, nach 10 Kontaktversuchen lagen 80 Prozent der später insgesamt erzielten Interviews vor. Rund sieben Prozent der Interviews kamen erst nach mehr als 15 Kontaktversuchen zustande. In Einzelfällen wurden Interviews sogar erst nach mehr als 30 Kontaktversuchen realisiert.

Abbildung 5 Anzahl der Kontaktversuche bis zum Interview: RCS-Studie zur Bundestagswahl, NAES 2000 und CES 1988 - 2000



Quelle: RCS-Studie „Kampagnendynamik 2005“, National Annenberg Election Study 2000 (vgl. Romer u. a. 2004: 63; Email-Mitteilung von Richard Johnston am 19.01.2005), Canadian Election Studies 1988 - 2000 (vgl. Northrup 1989: 10, 1994: 16, 1998: 13, 2003: 11)

Tabelle 1 zeigt, von welchen Faktoren es abhing, ob ein Interview rasch oder erst nach zahlreichen vergeblichen Kontaktversuchen zustande kam. Aufgrund der erwähnten Trunkierung der verfügbaren Feldzeit für späte Replikate durch den Wahltermin ist klar, dass die Zahl der überhaupt für Interviews verfügbaren Tage eine wesentliche Rahmenbedingung darstellt, die kontrolliert werden muss, auch wenn sie selbst keinen substanziellen Erkenntniswert besitzt. Bei einer breiter angelegten explorativen Analyse konnten vier unabhängige Variablen identifiziert werden, die sich auf die Zahl der Kontaktversuche

auswirkten, die für ein Interview nötig waren; sie betreffen ausnahmslos persönliche Merkmale der Befragten.¹⁴ Eine höhere Anzahl von Kontaktversuchen war demnach nötig, um Interviews mit jüngeren Personen, mit Hochgebildeten, mit Vollzeitbeschäftigten und mit Personen in kleinen Haushalten, insbesondere Ein-Personen-Haushalten zu realisieren. Dies sind durchweg Merkmale, die eine geringere Erreichbarkeit für telefonische Kontaktversuche in Privathaushalten indizieren und bedingen, dass im Schnitt intensivere Kontakthanstrengungen erforderlich sind, um mit solchen Personen Befragungen durchzuführen.¹⁵

**Tabelle 1 Zahl der Kontaktversuche bis zum erfolgreichen Interview:
Determinanten**

	Standardisierter Regressions- koeffizient	t-Wert	p
Zahl der in Feldzeit verfügbaren Tage (1-21)	.27	17.23	.000
Alter (in Jahren)	-.09	-5.36	.000
Abitur (dummy)	.05	3.31	.001
Vollzeit beschäftigt (dummy)	.11	7.00	.000
Haushaltsgröße (Zahl der Personen)	-.04	-2.26	.024
N = 3.559			

Modellgüte $R^2=10.7$ Prozent; dabei erklären die Kontrollvariablen eine zusätzliche Varianz von 3.1 Prozent gegenüber einem Modell, in das nur die Feldzeit eingeht.

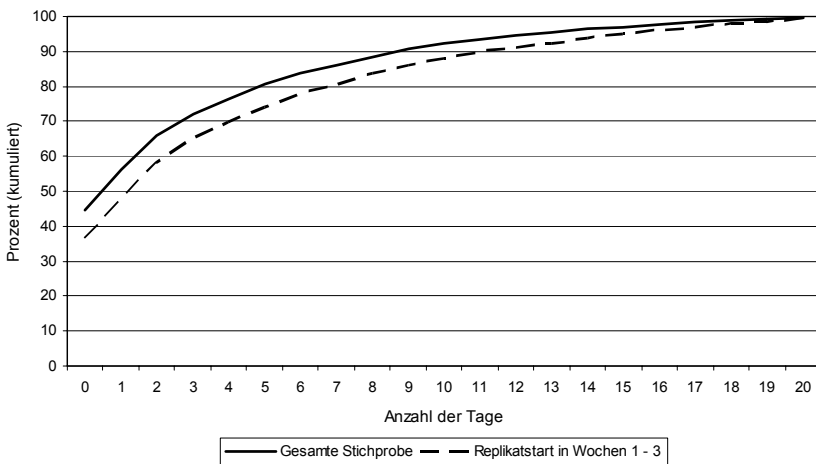
Daraus folgt unmittelbar, dass es mitunter sehr lange dauerte, bis Interviews realisiert werden konnten. Analog zu Abbildung 5 gibt Abbildung 6 wieder, wie viele Tage nach Replikatstart die vorhandenen Interviews durchgeführt wurden. Knapp 45 Prozent aller Interviews wurden ohne Verzögerung sofort am Tag der Freigabe ihres Replikats realisiert

14 Als nicht signifikante Prädiktoren der Zahl der Kontaktversuche bis zum erfolgreichen Interview erwiesen sich das Geschlecht der Befragten, ihre Kirchengangshäufigkeit, ihre Gewerkschaftsmitgliedschaft, die Größe ihres Wohnortes, die Wohnregion (alte oder neue Bundesländer), das politische Interesse, die Stärke der Parteibindung sowie die Zahl der Telefonanschlüsse im Haushalt und die Wiederbefragungsbereitschaft als persönliche Merkmale sowie der Wochentag des Replikatstarts als Merkmal der Erhebung.

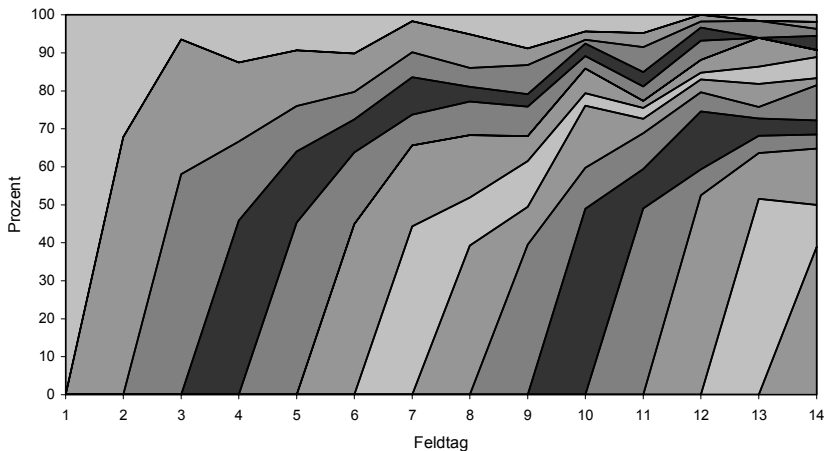
15 Die Zahl der für ein Interview der RCS-Studie nötigen Kontaktversuche ist positiv mit der Zahl der Kontaktversuche bei der 2. Welle und (leicht) negativ mit dem Vorliegen eines Interviews der 2. Welle korreliert (Pearsons Korrelationskoeffizient $r = .14$ bzw. $r = -.09$, jeweils signifikant mit $p < .001$ bei einem zweiseitigen Test).

(durchgezogene Linie), 56 Prozent innerhalb der ersten beiden Tage, an denen auch die intensivsten Kontakthanstrengungen unternommen wurden. 84 Prozent aller vorliegenden Interviews entstammen der ersten Feldwoche, aus der dritten Feldwoche kommen knapp fünf Prozent. Dabei ist aber zu bedenken, dass dieser „Erfolgswert“ für die lange Feldzeit artifiziell gedämpft ist, denn schon nach der Hälfte der insgesamt sechswöchigen Erhebungsperiode der Studie gingen nur noch Replikate ins Feld, die wegen des herannahenden Wahltermins nicht mehr volle 21 Tage lang ausgeschöpft werden konnten. Das verdeutlicht der Vergleich mit der gestrichelten Linie, welche den Feldverlauf nur für diejenigen Replikate wiedergibt, die in den ersten drei Wochen freigegeben wurden und deshalb noch vor der Wahl vollständig abgearbeitet werden konnten. Von diesen Interviews stammen deutlich höhere Anteile aus späteren Phasen der jeweiligen Öffnungsperioden. Sechs Prozent wurden noch in der dritten Woche realisiert, über 22 Prozent in der zweiten Woche. Lediglich ein gutes Drittel dieser Befragungen wurde sofort am ersten Tag durchgeführt, die 50-Prozent-Marke wurde erst am dritten Tag erreicht.

Abbildung 6 Tage zwischen Replikatfreigabe und Interview



Quelle: RCS-Studie „Kampagnendynamik 2005“

Abbildung 7 Anteile der Replikate an den täglichen Interviews

Quelle: RCS-Studie „Kampagnendynamik 2005“

Abbildung 7 präsentiert eine Art Resümee der bisherigen Beobachtungen. Abbildung 1 gab einen Eindruck von der idealtypischen Datenstruktur einer RCS-Umfrage, diese Darstellung zeigt nun die tatsächliche Datenstruktur der in den ersten 14 Tagen der Erhebungsperiode unserer RCS-Studie realisierten Interviews. Wegen der Fluktuation der täglichen Fallzahlen basiert diese Darstellung im Interesse der Übersichtlichkeit allerdings auf relativen Häufigkeiten innerhalb einzelner Feldtage. Offenkundig war bereits am fünften Feldtag eine gute Mischung von Interviews aus mehreren Replikaten erreicht, deren dominante Struktur sich später nur noch wenig änderte. Zu fragen ist nun, ob diese Daten auch tatsächlich die Erwartungen erfüllen, denen RCS-Studien gerecht werden sollen.

4 Zur Qualität der RCS-Daten

4.1 Immanente Qualität

Wie steht es um die Qualität unserer Daten gemäß der immanenten Kriterien des RCS-Designs? Das Kernanliegen einer solchen Studie besteht darin, eine Stichprobe zu realisieren, für deren Mitglieder nicht nur die Inklusion selbst, sondern auch der Tag, an dem sie interviewt werden, per Zufall bestimmt wird. Wenn es gelungen ist, dies adäquat um-

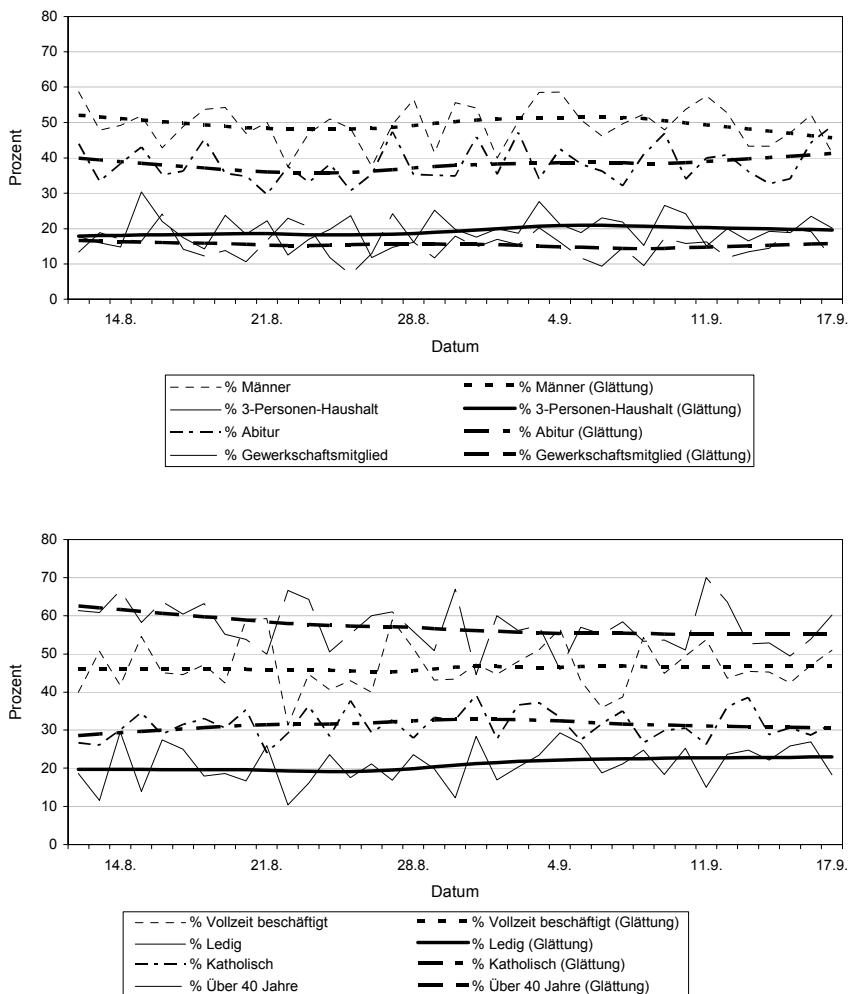
zusetzen, dann kann davon ausgegangen werden, dass die Verteilungen demographischer Variablen und anderer stabiler Merkmalsdimensionen im Feldverlauf nicht variieren. Ob dies der Fall ist, wird im Folgenden getestet. Für RCS-Daten bieten sich in besonderer Weise graphische Verfahren der Datenanalyse an (Romer u. a. 2004; Brady & Johnston 2006; Clarke u. a. 2004, 2006). Trendplots von Skalenmittelwerten oder Anteilen interessierender Merkmale empfehlen sich stets als Einstieg, um einen ersten Eindruck von der dynamischen Entwicklung theoretisch interessanter Aspekte der öffentlichen Meinung im Erhebungszeitraum zu gewinnen. Graphische Analysen sind daher ein häufig benutztes Instrument in RCS-Projekten, das die formale Modellierung von Variablenzusammenhängen nicht nur vorbereitet, sondern auch sinnvoll ergänzt. Wenn man eine solche Analyse auf der Ebene der einzelnen Stichprobentage durchführen will, ist man allerdings mit dem Problem konfrontiert, dass die einzelnen Tagesstichproben jeweils relativ klein und daher zwangsläufig mit einem recht großen Zufallsfehler behaftet sind. Dem kann durch Verfahren der Datenglättung entgegengewirkt werden. Die meisten vorliegenden Analysen von RCS-Daten verwenden hierfür einfache Verfahren gleitender Durchschnitte. Anspruchsvollere und solchen Herangehensweisen überlegene Techniken der Punktschätzung mittels Verfahren der nicht-parametrischen Regression (Fox 2000; Elff 2006) wurden bislang nur selten eingesetzt (siehe jedoch Clarke u. a. 2004, 2006). Wir werden im Folgenden auf den Scatterplot-Smoother LOWESS zurückgreifen.¹⁶ Die Analysen beginnen jeweils am fünften Tag der Feldperiode; die in der Anlaufphase der Erhebung gewonnenen Interviews können – wie oben diskutiert – designbedingt nicht als Zufallsstichproben gelten und werden deshalb bei dieser Analyse nicht berücksichtigt.

Wenn das Erhebungsverfahren das gewünschte Ergebnis – eine über Zeit völlig gleichmäßige Stichprobe – geliefert hat, bei der die Fälle jedes einzelnen Tages in sich eine Zufallsstichprobe darstellen, dann müssen unveränderliche demographische Merkmale in allen Tagesstichproben gleich verteilt sein. Systematische Trends dürfen nicht erkennbar

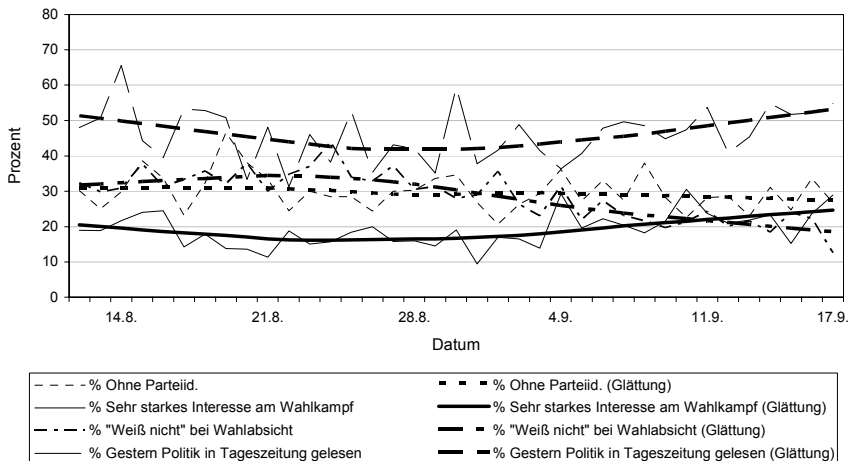
16 LOWESS steht für „*LO*cally *WE*ighted Scatterplot Smoother“. Dieses Glättungsverfahren berechnet für jeden einzelnen Datenpunkt einer Punktwolke eine lokale Regression, in die nur die „nächsten Nachbarn“ dieses Punktes eingehen – im hier vorliegenden Fall für jeden Tag also die Werte der Befragten des jeweils gleichen sowie der benachbarten Tage. Für jeden Datenpunkt wird aus den beobachteten Punkten, welche diesen umgeben, ein Schätzwert ermittelt. Eine Gewichtungsfunktion sorgt dabei dafür, dass der Einfluss dieser Punkte mit zunehmender Distanz abnimmt; weiter entfernte Punkte determinieren dadurch den Schätzwert für einen Punkt schwächer als näher liegende. Dabei können verschiedene so genannte Bandbreiten gewählt werden. Allen in diesem Artikel präsentierten Analysen liegt eine Bandbreite von .5 zugrunde, d. h. für die Schätzung jedes Punktes wurde – mit abnehmender Gewichtung – die Hälfte aller Beobachtungen berücksichtigt. Die Verbindung dieser Schätzwerte ergibt am Ende die geglättete Trendlinie (vgl. Cleveland 1994: 168-180; Schnell 1994: 109-113; Fox 2000: 19-50).

werden, lediglich zufallsbedingte Fluktuationen sind zulässig. Demographische Merkmale dürfen also nicht mit den Zeitpunkten korreliert sein, zu denen Befragte interviewt wurden. Abbildung 8 belegt für einige exemplarisch ausgewählte Merkmale, dass dies in unserer Studie in der Tat weitgehend der Fall ist. Die Anteile der Männer, der Personen, die in Drei-Personen-Haushalten leben, der Abiturienten und der Gewerkschaftsmitglieder weisen zwar geringfügige Schwankungen um die jeweiligen Mittelwerte auf, d. h. die Feldverläufe waren nicht immer vollkommen eben, aber eine Tendenz nach oben oder unten lässt sich nicht ausmachen.¹⁷ Dasselbe gilt auch für die Anteile der Vollzeit-Erwerbstätigen und der Katholiken. Etwas anders verhält es sich bei den Ledigen – ihr Anteil scheint über die Zeit leicht zuzunehmen. Der Anteil der Befragten, die 40 Jahre und älter sind, nimmt demgegenüber leicht ab. Etwa drei Wochen nach Beobachtungsbeginn tritt dann jeweils eine Stabilisierung ein. Wir haben oben gesehen, dass Personen, die erst nach vielen Kontaktversuchen befragt werden konnten, im Schnitt jünger sowie häufiger voll erwerbstätig waren. Abbildung 8 deutet darauf hin, dass sich die Stichprobenstruktur bezüglich dieser Merkmale erst dann stabilisiert hat, als die ersten Replikate abgeschlossen wurden, so dass die Replikatekomposition der Tagesstichproben keinen weiteren Veränderungen mehr unterworfen war. Insgesamt sind aber auch diese Schwankungen gering.

17 Eine inferenzstatistische Überprüfung bestätigt den Augenschein: Für alle möglichen Paare von Tagesstichproben – bei insgesamt 37 Tagesstichproben sind dies 666 – haben wir den Anteil der Männer und der Abiturienten mittels logistischer sowie das mittlere Alter mittels linearer Regressionen in den beiden jeweils betrachteten Tagesstichproben verglichen. In die Regressionsmodelle haben wir jeweils neben einer Konstanten nur eine einzige Dummy-Variable, die die Zugehörigkeit zu *einer* der beiden einbezogenen Tagesstichproben anzeigt, einbezogen. Für den Anteil der Männer erreichte der Koeffizient dieser Variable in 44 Fällen (6,6 Prozent) statistische Signifikanz ($p < .05$). Für den Anteil der Abiturienten bzw. das mittlere Alter waren es 54 (8,1 Prozent) bzw. 34 (5,1 Prozent). Da bei einem Signifikanzniveau von fünf Prozent 33 signifikante von insgesamt 666 Koeffizienten zu erwarten sind, deuten diese Werte auf fast ausschließlich zufällige Schwankungen der sozialstrukturellen Merkmale hin.

Abbildung 8 Verteilungsstabilität demographischer Merkmale

Quelle: RCS-Studie „Kampagnendynamik 2005“

Abbildung 9 Verteilungsentwicklung politischer Orientierungen

Quelle: RCS-Studie „Kampagnendynamik 2005“

Systematische Veränderungen der demographischen Zusammensetzung der Stichprobe von Erhebungstag zu Erhebungstag wären ein Hinweis auf mangelhafte Implementation des RCS-Designs. Abbildung 8 deutet darauf hin, dass solche Probleme nicht existieren. Anders verhält es sich bei den in Abbildung 9 wiedergegebenen politischen Orientierungen. Hier ist keine Stabilität über Zeit zu erwarten, jedenfalls dann nicht, wenn es sich nicht um tief in der Persönlichkeit verwurzelte Grundhaltungen handelt. Im Gegenteil – Wahrnehmungen und Einstellungen sollten sich ändern, denn die Erwartung, dass sich solche Orientierungen im Verlaufe von Wahlkämpfen wandeln, ist eine zentrale Prämisse und die Rechtfertigung der Durchführung von RCS-Studien. Solche Dynamiken sichtbar zu machen, ist ja gerade ihr Zweck. Abbildung 9 illustriert diese Überlegungen am Beispiel von vier unterschiedlichen politischen Orientierungen. Abgetragen ist zunächst der Anteil derjenigen Wähler, die sich nicht mit einer politischen Partei identifizieren. Die orthodoxe Theorie der Parteidentifikation ging davon aus, dass solche Orientierungen außerordentlich stabil sind und sich unter normalen politischen Umständen kaum jemals ändern (Campbell u. a. 1960). Jüngere Deutungen halten Parteibindungen zwar für durchaus wandelbar; aber auch sie interpretieren sie als eher stabile Komponenten der politischen Persönlichkeit (Fiorina 1981; Schmitt-Beck & Weick 2001; Schmitt-Beck u. a. 2006). Abbildung 9 deutet darauf hin, dass sich der Anteil der parteipolitisch Unabhängi-

gen (und damit implizit auch der parteigebundenen Wähler) in der Tat während des Wahlkampfes zur Bundestagswahl 2005 praktisch nicht verändert hat. Insoweit bekräftigt dieser Befund den aus der Analyse der demographischen Variablen gezogenen Schluss zur immanenten Qualität der Daten.

Im Vergleich dazu waren die anderen drei politischen Orientierungen relativ beweglich. Sehr deutlich wird erkennbar, wie der Wahlkampf den Wählern geholfen hat, Präferenzen zu formieren. Der Anteil derjenigen Befragten, die zum jeweiligen Befragungszeitpunkt noch nicht wussten, wie sie wählen sollten, halbierte sich in den letzten vier Wochen vor der Wahl kontinuierlich von etwa 35 auf unter 20 Prozent. Im selben Zeitraum wuchs das Interesse am Wahlkampf. Korrespondierend dazu nahm der Anteil der Befragten zu, die angaben, sich am Vortag durch Lektüre einer Tageszeitung politisch informiert zu haben. Der nächste Abschnitt wird belegen, dass auch die Parteipräferenzen selbst deutlichen Veränderungsprozessen unterworfen waren.

4.2 Die Entwicklung von Parteipräferenzen: die RCS-Daten im Vergleich mit anderen Datenquellen

Im Mittelpunkt unseres Projektes stehen letztlich die Wahlabsichten und ihre Entwicklung während des Wahlkampfes bis hin zu ihrer letztendlichen Manifestation in Form des Ergebnisses der Bundestagswahl am 18. September 2005, das am Ende des Wahlkampfes verzeichnet wurde und zumindest partiell als Reaktion auf Prozesse der Präferenzbildung und -änderung in dessen Verlauf verstanden werden kann. Der RCS-Studie zufolge hat es im Verlauf des Wahlkampfes einige deutliche Verschiebungen in den Parteipräferenzen gegeben. Diese vergleichen wir im Folgenden mit anderen Messungen der Unterstützung der Parteien und ihrer Entwicklung. Naheliegend ist selbstverständlich zunächst der Vergleich mit dem Wahlergebnis selbst. Allerdings interessiert hier keineswegs nur die statische, auf den 18. September fokussierte Perspektive, sondern ebenso die dynamische Perspektive auf Entwicklungen im Vorfeld. Wie stellt sich diese Entwicklung der Unterstützung der Parteien über den gesamten Wahlkampf hinweg in den RCS-Daten dar?

Ein vergleichbares Anliegen verfolgen auch die regelmäßig zum Zwecke der Medienpublikation produzierten Versuche von Meinungsforschungsinstituten, Stimmenverteilungen für eine hypothetische Bundestagswahl zu einem bestimmten, vor der tatsächlichen Wahl liegenden Termin zu ermitteln. Die „demoskopischen“ Institute achten zwar normalerweise darauf, dabei den Begriff „Wahlprognose“ zu vermeiden, und weichen stattdessen auf weniger verfängliche Begriffe wie „Projektion“ oder einfach „Sonntagsfrage“ aus (vgl. ZDF 2006). Aber sie können dadurch kaum verhindern, dass die in vielen Medien publizierten Fieberkurven der Wählermeinung in der öffentlichen Wahrnehmung und Diskussi-

on häufig wie Vorhersagen des tatsächlichen Wahlergebnisses der kommenden Bundestagswahl verarbeitet werden und sich dadurch möglicherweise auch auf das Stimmverhalten der Wähler auswirken (Brettschneider 1992; Schmitt-Beck 1996). Das ist auch der Hintergrund der Debatte über die Qualität von Medienumfragen, welche – wie schon in vergleichbaren Situationen bei früheren Wahlen – sofort nach der Bundestagswahl 2005 einsetzte, weil die vor der Bundestagswahl publizierten Zahlen vom Wahlergebnis in eklatanter Weise abwichen (Hofmann-Göttig 2005). So schienen wieder einmal die „Demoskopien“ ein „Debakel“, wenn nicht gar ein „Desaster“ zu erleben und wurden als die „eigentlichen Wahlverlierer“ apostrophiert (vgl. Spiegel 2005, NZZ 2005, FTD 2005), während sich die SPD am Ende ihres Wahlkampfes als „Umfragesieger-Besieger“ (Das Parlament 2005) brüsten konnte.

Eine weitere Ausdrucksform der Entwicklung der öffentlichen Meinung gegenüber den Parteien im Verlauf des Wahlkampfes stellen die Kurse von Wahlbörsen dar. Eine Wahlbörse funktioniert wie ein Aktienmarkt – Aktien werden zu aktuellen Marktpreisen gekauft und verkauft. Im Falle von Wahlbörsen handelt es sich dabei um „Parteiaktien“. Da dem Besitzer einer Parteiaktie nach der Bundestagswahl ein Betrag in Höhe des Zweitstimmenergebnisses der jeweiligen Partei ausgezahlt wird, spiegeln die Kurse zu jedem Zeitpunkt die durchschnittlichen Erwartungen der Händler hinsichtlich des Wahlausgangs wider. Der aktuelle Kurs einer Aktie entspricht jeweils dem Preis, zu dem die letzte Transaktion getätigt wurde. Die Kurse können wie Stimmenanteile interpretiert werden. Anders als Befragungspersonen in Umfragen bringen die Teilnehmer von Wahlbörsen allerdings keine momentanen persönlichen Präferenzen zum Ausdruck, sondern ihre Vermutungen über den Ausgang der Wahl und damit letztlich über das voraussichtliche Verhalten aller anderen Wähler. Im Gegensatz zu Meinungsumfragen werden Wahlbörsen von ihren Betreibern daher als echte Prognoseinstrumente für das Wahlergebnis propagiert, die sich gleichsam der kollektiven „Weisheit des Marktes“ bedienen und dadurch sensibel auf aktuelle politische Entwicklungen reagieren (Brüggelambert 1999; Huber 2002). Wir berücksichtigen bei unserer Analyse Daten der *Wahlbörse* sowie der *Wahlstreet*.¹⁸ Um auf eine tageweise Basis zu kommen, betrachten wir die gemittelten täglichen Schlusskurse der beiden Börsen, d. h. für jede Partei den (mittleren) Preis der letzten Transaktion des Tages.

18 Beide Wahlbörsen wurden von ECCE TERRAM veranstaltet – die *Wahlbörse* im Auftrag von Financial Times Deutschland sowie T-Online, die *Wahlstreet* im Auftrag der ZEIT sowie des Handelsblatt (vgl. <http://wahlboerse.t-online.de>; <http://www.wahlstreet.de>). Unser Dank gilt Frank Simon von ECCE TERRAM, der uns die Daten freundlicherweise zur Verfügung gestellt hat.

Diese drei Datenquellen – RCS-Studie, Medienumfragen¹⁹ sowie Wahlbörsen – wollen wir daher mit Blick auf den 18. September selbst, aber auch mit Blick auf die Dynamik im Vorfeld des 18. September vergleichen. Anders als die in Abbildung 8 und 9 wiedergegebenen Befunde berücksichtigt die nachfolgende Analyse nur diejenigen Befragten unserer RCS-Studie, die angaben, zumindest „vielleicht“ an der Bundestagswahl teilnehmen zu wollen, und bei der Frage nach dem beabsichtigten Wahlverhalten eine Partei nannten.²⁰ Um die Zufallsschwankungen der tagesgenau gemessenen und damit nur auf relativ geringen Fallzahlen beruhenden Verteilungen zu dämpfen, werden die Daten wieder einer LOWESS-Glättung unterzogen.

Bekanntlich waren bei der Bundestagswahl 2005 alle Formen der Datensammlung mit einem schwierigen Umfeld konfrontiert, wie auch die einzelnen Grafiken in Abbildung 10 zeigen: Vergleicht man die Endpunkte der verschiedenen Kurven mit dem tatsächlichen Wahlergebnis vom 18. September 2005, so liefert letztlich keine Datenquelle eine exakte Entsprechung zum Wahlergebnis vom 18. September 2005 über alle Parteien hinweg. Insbesondere das schlechte Abschneiden der Union deutet sich in keiner Weise an.

Welches Bild zeichnen die verschiedenen Datenquellen aber für die Dynamik im Vorfeld der Wahl?²¹ Auffällig ist bei dieser Betrachtung zunächst, wie eng die von Medienumfragen und Wahlbörsen ermittelten Parteienteile stets beieinander lagen. Möglicherweise

19 Die Zeitreihen der Medienumfragen wurden wie folgt gebildet: Es wurden alle veröffentlichten „Sonntagsfragen“ bzw. „Projektionen“ von Emnid, Forsa, der Forschungsgruppe Wahlen, Infratest dimap sowie dem Institut für Demoskopie Allensbach im Zeitraum vom 13. August bis zum 18. September 2005 gesammelt (siehe auch <http://wahlrecht.de/umfragen/archiv/2005.htm>). Um daraus tageweise organisierte Zeitreihen zu bilden, wurden den einzelnen Tagen die Parteienteile aus der jeweils zu diesem Zeitpunkt aktuellsten Umfrage zugewiesen. Die Zeitreihen zeigen also unabhängig vom Institut für jeden Tag die publizierten Verteilungen der zuletzt veröffentlichten „Sonntagsfrage“ an. Wurden an einem Tag zwei neue Umfragen unterschiedlicher Institute veröffentlicht, wurden daraus Mittelwerte gebildet.

20 Für den Vergleich mit Medienumfragen und Wahlbörsen ist eine möglichst realitätsnahe Schätzung der täglichen (Rand-)Verteilungen dieser Frage wichtig. Da die RCS-Daten die bei Telefonumfragen oft zu beobachtende Überrepräsentation hoch gebildeter Befragter aufweisen (Glasgow & Bühler 1987; Mosmann 1998), wurde bei den Berechnungen nach Bildung korrigiert. Das verwendete Gewicht passt die einzelnen Tagesstichproben an die Bildungsverteilung des ALLBUS 2004 an.

21 Diese vergleichende Darstellung kann dabei nicht im selben Sinne wie der Vergleich mit dem Wahlergebnis (im Sinne der Kriteriumsvalidierung) interpretiert werden, weil die Konstruktvalidität der „Projektionen“ abgesehen vom unmittelbaren zeitlichen Umfeld der Wahl eine unbekannte Größe ist. Es steht dann nämlich kein Außenkriterium als Vergleichsmaßstab zur Verfügung und die Institute legen die Verfahren, über welche sie ihre Rohdaten in „Projektionen“ transformieren, nicht offen (Wüst 2003). Ähnliches gilt auch für die Wahlbörsen. Der Vergleich unserer RCS-Studie mit diesen Daten stellt somit letztlich unterschiedliche Repräsentationen der öffentlichen Meinung einander gegenüber.

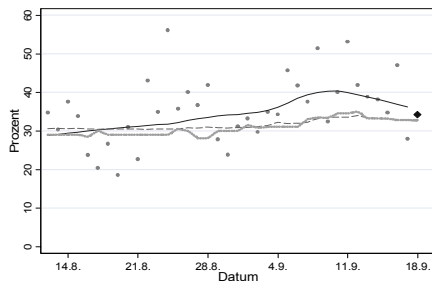
fließen publizierte Umfrageergebnisse als wesentliche Information in die Kaufentscheidungen der Händler an den Wahlbörsen ein, so dass deren Kurse faktisch gar keine unabhängigen Messungen darstellen.

Vergleicht man die geglätteten RCS-Werte über den gesamten hier betrachteten Zeitraum hinweg mit Werten aus Medienumfragen bzw. Wahlbörsen, zeigen sich viele Gemeinsamkeiten, aber auch einige interessante Unterschiede. Für die SPD etwa zeichnen sowohl die Medienumfragen als auch die Wahlbörsen zunächst ein träges Bild um die 30 Prozent; erst nach dem TV-Duell vom 4. September steigen die Werte der Sozialdemokraten bis auf 32 oder 33 Prozent an. Die RCS-Daten verzeichnen hingegen schon früher ein leichtes Ansteigen des SPD-Anteils, das sich nach dem TV-Duell deutlich beschleunigt. Überdies – und das ist das eigentlich Überraschende – legen die RCS-Daten den Schluss nahe, dass die Unterstützung der SPD zwar nach dem TV-Duell einen Höhepunkt erreichte, in den Tagen unmittelbar vor der Wahl aber wieder etwas schrumpfte. In der sport-affinen Sprache der Wahlkampfberichterstattung ausgedrückt, befand sich die SPD am Ende demnach gar nicht mehr in einer Aufholjagd – vielmehr ging ihr gegen Ende des Wahlkampfes eher schon wieder die Puste aus.

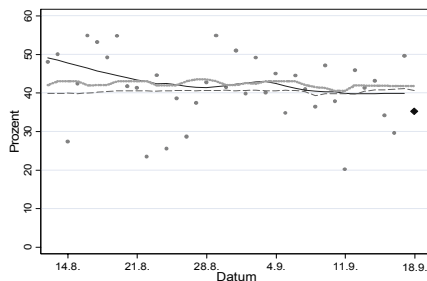
Für die Union stellt sich das Bild deutlich weniger bewegt dar, und zwar unabhängig von der Datenquelle. Sowohl für die RCS-Studie als auch für Medienumfragen und Wahlbörsen werden konstant Werte im Bereich um 40 Prozent für die Union registriert, die drei Zeitreihen überlappen sich stark. Den Niedergang der Union vor dem Wahltag, der schließlich zum Wahlergebnis von 35.2 Prozent führte, bildet keine der Datenreihen adäquat ab. Allerdings fallen die ungeglätteten Datenpunkte der RCS-Studie in den letzten Tagen vor der Wahl auffallend niedrig aus; lediglich am allerletzten Tag vor der Wahl wird der Stimmenanteil der Union – möglicherweise zufallsbedingt – wieder klar überschätzt. Bei der Glättung führen diese beiden gegensätzlichen Entwicklungen zu einer Nivellierung, aus der dann kein Trend mehr ablesbar ist. An den Schwierigkeiten, solche sich mutmaßlich sehr schnell vollziehenden und im Volumen nicht übermäßig starken, in den machtpolitischen Konsequenzen aber entscheidenden Bewegungen kurz vor der Wahl adäquat abzubilden, zeigen sich designbedingte Grenzen von RCS-Studien, die allerdings nicht systematischer, sondern forschungspragmatischer Natur sind. Mit mehr Befragten in den letzten Tagen der Erhebung hätte sich der Zufallsfehler reduzieren lassen und es wären genauere Messungen zustande gekommen – allerdings nur um den Preis einer sehr deutlichen Steigerung der Erhebungskosten.

Abbildung 10 **Entwicklungen der Parteistärken im Vorfeld der Bundestagswahl 2005: RCS-Studie im Vergleich zu Mediumfragen und Wahlbörsen**

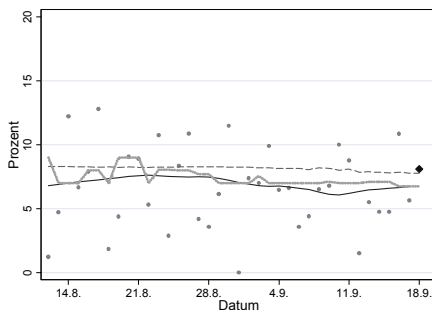
a) SPD



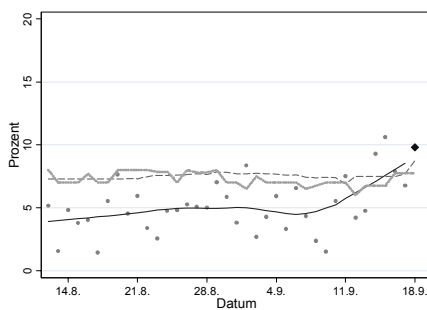
b) Union



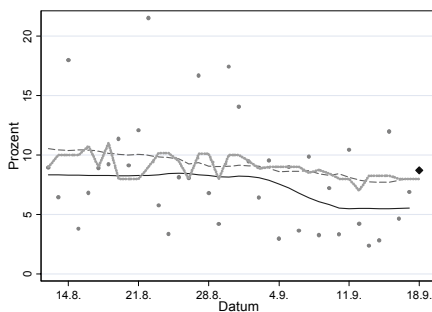
c) Grüne



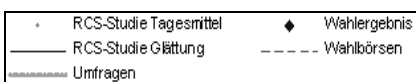
d) FDP



e) Linkspartei



Legende



Quelle: RCS-Studie „Kampagnendynamik 2005“, publizierte Umfragen von EMNID, Forsa, der Forschungsgruppe Wahlen, Infratest dimap und des Instituts für Demoskopie Allensbach sowie tägliche Schlusskurse der Internet-Wahlbörsen „Wahlstreet“ und „Wahlbörse“

Die Grünen liegen in den Wahlbörsen konstant bei etwa acht Prozent, ähnlich auch in den Medienumfragen. Die RCS-Daten deuten hingegen auf etwas stärkere Schwankungen hin und zeigen gerade am Wahlkampfsende (leicht) wachsenden Zuspruch für die Grünen. Möglicherweise ist das ein Effekt strategischen Wahlverhaltens in Reaktion auf die in der letzten Phase des Wahlkampfes intensivierte, auf rot-grün-orientierte SPD-Anhänger zielende Zweitstimmenkampagne der Grünen. Dies könnte auch den gleichzeitigen leichten Schwund bei den SPD-Anteilen plausibel machen. Besonders interessant ist der Befund für die FDP. Wiederum zeichnen Wahlbörsen und Medienumfragen ein ziemlich konstantes Bild. An den RCS-Daten stechen jedoch zwei Dinge klar hervor: dieser Studie zufolge lag die FDP lange Zeit nur im Bereich um fünf Prozent (und damit deutlich unter den übrigen Werten des Prognosemarktes), um dann aber einen fulminanten Schlusspunkt hinzulegen: In der letzten Wahlkampfwoche erst steigt die FDP auf ihr annähernd zweistelliges Wahlergebnis, was die LOWESS-Anpassungskurve auch klar sichtbar macht.²² Auch hierbei hat es sich vermutlich um ein Phänomen strategischen Wählens gehandelt, dieses Mal jedoch im bürgerlichen Lager. Die Werte der Linkspartei haben sich im Verlauf des Wahlkampfes kontinuierlich verschlechtert – diesen Schluss legen *alle* Trendlinien nahe. Dieses Absinken wird in den RCS-Daten allerdings noch markanter (und letztlich wohl zu markant) sichtbar als in Medienumfragen und Wahlbörsen.

Insgesamt legt diese vergleichende Analyse den Schluss nahe, dass das „Debakel der Demoskopen“ bei der Bundestagswahl 2005 zumindest zum Teil auf strategische Wahlentscheidungen zurückzuführen ist, die erst in den letzten Vorwahltagen getroffen wurden und die großen Parteien zugunsten ihrer jeweiligen tatsächlichen oder angestrebten Koalitionspartner Stimmen gekostet haben. Besonders ausgeprägt scheinen solche Wanderungsbewegungen im „schwarz-gelben Lager“ stattgefunden zu haben, dessen Ergebnis die „Prognosen“ auch besonders schlecht vorhersagten. Während diese Interpretationen nur vorläufiger Natur sein können und durch weitere Analysen erhärtet werden müssen, vermitteln die Befunde der in diesem Kapitel dargestellten Analysen auf jeden Fall einen deutlichen Eindruck vom Nutzen tagesgenauer Beobachtung öffentlicher Meinung im Verlauf von Wahlkampagnen mit dem Instrument der RCS-Umfrage.

22 Für die FDP zeigt sich in der *Wahlstreet* ein interessantes Phänomen: Am Wahltag schießt der FDP-Kurs ab ungefähr 16 Uhr auf annähernd 10 Prozent nach oben. Die Vermutung liegt nahe, dass hier ein Händler Profite aus Insiderwissen erzielen wollte, denn gegen 16 Uhr lagen am Wahltag „gut informierten Kreisen“ die ersten Zahlen aus den Wahltagsbefragungen vor.

5 Fazit

Im vorliegenden Beitrag ging es zunächst darum, die zentralen Ideen eines relativ neuen Erhebungsverfahrens zur Erfassung kurzfristiger Prozesse des Wandels der öffentlichen Meinung vorzustellen: der *Rolling Cross-Section*-Umfrage. Im nächsten Schritt wurde diskutiert, wie dieses Design bei der ersten deutschen RCS-Studie umgesetzt wurde, die anlässlich der Bundestagswahl 2005 durchgeführt wurde. Es wurde deutlich, dass die praktische Realisierung einer solchen Umfrage Primärforscher und Feldinstitute in verschiedener Hinsicht vor beträchtliche Herausforderungen stellt und große Ansprüche an die Qualität der Feldarbeit und eine kontinuierliche Studienbetreuung impliziert. Die im dritten Teil des Artikels vorgestellten Befunde aus unserer Studie können jedoch als Beleg dienen, dass sich der Aufwand (und der damit verbundene finanzielle Einsatz) lohnt. An diesen wenigen exemplarischen Befunden, die das analytische Potenzial der Daten bei weitem nicht ausschöpfen, wird unschwer ersichtlich, dass RCS-Daten neue Einblicke in Wahlkämpfe und andere dynamische Phänomene der öffentlichen Meinung ermöglichen.

Vor dem Hintergrund zunehmender Volatilität des politischen Verhaltens sind einfache Querschnittsstudien, auch wenn sie in monatlichen oder gar wöchentlichen Zeitabständen wiederholt werden, aber auch Panel-Studien immer weniger imstande, die Dynamik der Meinungsbildung und Präferenzformierung nachzuzeichnen. Die berichteten Ergebnisse deuten auf erhebliche Fluktuationen in den Wahlabsichten hin, die sich in den letzten Tagen unmittelbar vor der Wahl stark intensivierten. Diese Prozesse dürften zumindest eine Teilantwort auf das Rätsel der schlechten Prognoseleistungen der für die Medien durchgeführten Meinungsumfragen bei dieser Wahl implizieren. RCS-Erhebungen sind besonders geeignet, um Phänomene des „late swing“ aufzudecken, welche noch unmittelbar vor Wahlen zu erheblichen Verschiebungen der Stimmenanteile von Parteien führen können. Lediglich was am Wahltag selbst geschah, entging bei der Bundestagswahl 2005 selbst diesem Instrument. Erste Analysen unserer Nachwahl-Welle deuten darauf hin, dass fast ein Zehntel der Wähler noch nach Ende der RCS-Erhebung eine Wahlentscheidung getroffen oder revidiert hat. Durch die Kombination der Vorwahl-RCS-Erhebung mit einer Nachwahl-Panel-Welle besteht jedoch eine ideale Voraussetzung, um auch diesen Prozessen analytisch auf die Spur zu kommen.

Literatur

- Aarts, K., van der Kolk, H. & Kamp, M. (1999). *Dutch Parliamentary Election Study 1998*. Documentation, Enschede: NIWI/Steinmetz Archive/Dutch Electoral Research Foundation (SKON).
- Allsop, D. & Weisberg, H. F. (1988). Measuring Change in Party Identification in an Election Campaign. *American Journal of Political Science*, 32, 996-1017.
- Asher, H. (1998). *Polling and the Public. What Every Citizen Should Know*, 4th edition. Washington, DC: CQ Press.
- Bartels, L. M. (1988). *Presidential Primaries and the Dynamics of Public Choice*. Princeton/NJ: Princeton University Press.
- Bartels, L. M. (1999). Panel Effects in the American National Election Studies. *Political Analysis*, 8, 1-20.
- Bartels, L. M. (2006). Three Virtues of Panel Data for the Analysis of Campaign Effects. In H. E. Brady & R. Johnston (Eds.), *Capturing Campaign Effects*, forthcoming. Ann Arbor: University of Michigan Press.
- Blais, A., Gidengil, E., Nadeau, R. & Nevitte, N. (2002). *Anatomy of a Liberal Victory: Making Sense of the Vote in the 2000 Canadian Election*. Peterborough: Broadview Press.
- Blakeman, B. W. (1995). Tracking Polls: How to Do Them. *Campaigns & Elections*, 16, 24-25.
- Brady, H. E. (2000). Contributions of Survey Research to Political Science. *PS – Political Science and Politics*, 33, 47-57.
- Brady, H. E. & Johnston, R. (2006). *Capturing Campaign Effects*. Ann Arbor: University of Michigan Press (i. E.).
- Brady, H. E., Johnston, R. & Sides, J. (2006). The Study of Political Campaigns. In H. E. Brady & R. Johnston (Eds.), *Capturing Campaign Effects*, forthcoming. Ann Arbor: University of Michigan Press.
- Brettschneider, F. (1992). Der taktische und rationale Wähler. Über den Einfluß von Wahlumfragen auf das Wählerverhalten bei den Bundestagswahlen 1983 und 1990. *Politische Vierteljahresschrift*, 33, 55-72.
- Brüggelambert, G. (1999). *Entscheidungsverhalten und Institutionen als Informationsträger: Erfahrungen mit Wahlbörsen*. Marburg: Metropolis.
- Campbell, A., Converse, P. E., Miller, W. E. & Stokes, D. E. (1960). *The American Voter*. New York: Wiley.
- Clarke, H. D., Sanders, D., Stewart, M. C. & Whiteley, P. (2004). *Political Choice in Britain*. Oxford: Oxford University Press.

- Clarke, H. D., Sanders, D., Whiteley, P. & Stewart, M. (2006). Campaign Dynamics and Campaign Effects in the 2001 British General Election. In A. McCutcheon (Eds.), *The Science of Pre-Election Polling*, forthcoming. Lincoln: University of Nebraska Press.
- Cleveland, W. S. (1994). *The Elements of Graphing Data, revised edition*. Murray Hill/NJ: At&T Bell Labs.
- Clymer, A. (2006). *NAES 04. National Annenberg Election Study*. Retrieved April 7, 2006, from <http://www.annenbergpublicpolicycenter.org/naes/>
- Das Parlament (2005). *Der „Umfragesieger-Besieger“*. Retrieved April 7, 2006, from <http://www.das-parlament.de/2005/36/Thema/003.html>
- Dichter Nebel (2005). *Spiegel*, 55.
- Dobrzynska, A., Blais, A. & Nadeau, R. (2003). Do the Media Have a Direct Impact on the Vote? The Case of the 1997 Canadian Election. *International Journal of Public Opinion Research*, 15, 27-43.
- Dörner, C. (1998). Do Media Matter? – Zur Reichweite des Agenda Settings. In C. Dörner & Klaudia Erhardt (Eds.), *Politische Meinungsbildung und Wahlverhalten. Analysen zum „Superwahljahr“ 1994* (pp. 119-149). Wiesbaden: Westdeutscher Verlag.
- Elff, M. (2006). Nichtparametrische und semiparametrische Regressionsverfahren. In J. Behnke, T. Gschwend, D. Schindler & K. Schnapp (Eds.), *Methoden der Politikwissenschaft*, forthcoming. Baden-Baden: Nomos.
- Erhardt, K. (1998). Die unentschlossenen Wähler als Motor zyklischer Wahlabsichtsverläufe. Eine Untersuchung am Beispiel der Rückgewinnung der Wähler durch die CDU im Frühjahr 1994. In C. Dörner & K. Erhardt (Eds.), *Politische Meinungsbildung und Wahlverhalten. Analysen zum "Superwahljahr" 1994* (pp. 15-118). Wiesbaden: Westdeutscher Verlag.
- Farrell, D. M. & Schmitt-Beck, R. (2002). *Do Political Campaigns Matter? Campaign Effects in Elections and Referendums*. London/New York: Routledge.
- Fiorina, M. P. (1981). *Retrospective Voting in American National Elections*. New Haven: Yale University Press.
- Fournier, P., Nadeau, R., Blais, A., Gidengil, E. & Nevitte, N. (2004). Time-of-voting decision and susceptibility to campaign effects. *Electoral Studies*, 23, 661-681.
- Fox, J. (2000). *Nonparametric Simple Regression: Smoothing Scatterplots*. Thousand Oaks: Sage.
- FTD (2005). *Gallup-Institut kritisiert deutsche Meinungsforscher*. Retrieved April 7, 2006, from <http://www.ftd.de/pw/de/23035.html>
- Gabler, S. & Häder, S. (1997). Überlegungen zu einem Stichprobendesign für Telefonumfragen in Deutschland. *ZUMA-Nachrichten*, 41, 7-18.

- Gidengil, E. & Dobrzynska, A. (2003). *Using a rolling cross-section design to model media effects: The case of leader evaluations in the 1997 Canadian elections*. Paper prepared for delivery at the Annual Meeting of the APSA, Philadelphia.
- Glagow, H. & Bühler, E. (1987). *C.A.T.I. – Das Antwortverhalten*. Sample Workshop Nr. 2.
- Güllner, M. (2000). Methodische Entwicklungen in der Empirischen Wahlforschung. In M. Klein, W. Jagodzinski, E. Mochmann & D. Ohr (Eds.), *50 Jahre Empirische Wahlforschung in Deutschland. Entwicklung, Befunde, Perspektiven, Daten* (pp. 564-583). Wiesbaden: Westdeutscher Verlag.
- Güllner, M., Dülmer, H., Klein, M., Ohr, D., Quandt, M., Rosar, U. & Klingemann, H. (2005). *Die Bundestagswahl 2002. Eine Untersuchung in Zeiten hoher politischer Dynamik*. Wiesbaden: VS-Verlag.
- Häder, S. & Gabler, S. (1998). Ein neues Stichprobendesign für telefonische Umfragen in Deutschland. In S. Häder, S. Gabler & J. H.P. Hoffmeyer-Zlotnik (Eds.), *Telefonstichproben in Deutschland* (pp. 69-88). Opladen: Westdeutscher Verlag.
- Hillygus, D. S. & Jackman, S. (2003). Voter Decision Making in Election 2000: Campaign Effects, Partisan Activation, and the Clinton Legacy. *American Journal of Political Science*, 47, 583-596.
- Hofmann-Göttig, J. (2005). *Umfragen zur Bundestagswahl vom 18. September 2005: Was taugte die „Sonntagsfrage“?* Retrieved April 7, 2006, from http://www.mbfj.rlp.de/Wir_ueber_uns/downloads/Umfragen_zur_Bundestagswahl_2005.pdf
- Holbrook, T. M. (1996). *Do Campaigns Matter?* Thousand Oaks: Sage.
- Huber, J. (2002). *Wahlbörsen: Preisbildung auf politischen Märkten zur Vorhersage von Wahlergebnissen*. Hamburg: Kovac.
- Johnston, R. (2001). Capturing Campaigns in National Election Studies. In E. Katz & Y. Warshel (Eds.), *Election Studies. What's Their Use?* (pp. 149-172). Boulder/Col.: Westview.
- Johnston, R., Blais, A., Brady, H. E. & Crête, J. (1992). *Letting the People Decide. Dynamics of a Canadian Election*. Stanford/CA: Stanford University Press.
- Johnston, R., Blais, A., Gidengil, E. & Nevitte, N. (1996). *The Challenge of Direct Democracy. The 1992 Canadian Referendum*. Montreal: McGill-Queen's University Press.
- Johnston, R. & Brady, H. E. (2002). The rolling cross-section design. *Electoral Studies*, 21, 283-295.
- Johnston, R., Hagen, M. G. & Jamieson, K. H. (2004). *The 2000 Presidential Election and the Foundations of Party Politics*. Cambridge/Mass: Cambridge University Press.
- Lazarsfeld, P. F., Berelson, B. & Gaudet, H. (1944). *The People's Choice. How the Voter Makes Up his Mind in a Presidential Campaign*. New York/London: Columbia University Press.
- Mosmann, H. (1998). *Das Antwortverhalten PAPI – CAPI – CATI*. Intra Workshop Nr. 6.

- Nevitte, N. & Blais, A., Gidengil, E. & Nadeau, R. (2000). *Unsteady State: The 1997 Canadian Federal Election*. Don Mills: Oxford University Press.
- Northrup, D. (1989). *The 1988 National Election Study. Technical Documentation*. Toronto: Institut for Social Research, York University. Retrieved April 7, 2006, from http://www.isr.yorku.ca/download/CES/CES88/Techdoc_88.pdf
- Northrup, D. (1994). *The 1993 Canadian Election Study. Technical Documentation*. Toronto: Institut for Social Research, York University. Retrieved April 7, 2006, from <http://www.isr.yorku.ca/download/CES/CES93/TECHDOC.PDF>
- Northrup, D. (1998). *The 1997 Canadian Election Survey. Technical Documentation*. Toronto: Institut for Social Research, York University. Retrieved April 7, 2006, from http://www.isr.yorku.ca/download/CES/CES97/TECH_ENG.PDF
- Northrup, D. (2003). *The 2000 Canadian Election Survey. Technical Documentation*. Toronto: Institut for Social Research, York University. Retrieved April 7, 2006, from http://www.isr.yorku.ca/download/CES/CES00/CES_TD_001.pdf
- NZZ (2005). *Deutschlands Wahlforscher als Verlierer*. Retrieved April 7, 2006, from http://www.nzz.ch/dossiers/Wahlen_Deutschland/resultate/2005/09/21/al/articleD5S9A.html
- Romer, D., Kenski, K., Waldman, P., Adasiewicz, C. & Jamieson, K. H. (2004). *Capturing Campaign Dynamics: The National Annenberg Election Survey. Design, Method, and Data*. New York/Oxford: Oxford University Press.
- Sanders, D. J., Whiteley, P. F., Clarke, H. D., Stewart, M. & Winters, K. *The British Election Study. 2005 Rolling Campaign Survey Data*. Retrieved April 7, 2006, from <http://www.essex.ac.uk/bes/2005/ircp.htm>
- Schmitt-Beck, R. (1996). Mass Media, the Electorate, and the Bandwagon. A Study of Communication Effects on Vote Choice in Germany. *International Journal of Public Opinion Research*, 8, 266-291.
- Schmitt-Beck, R. (2005). *Kampagnen-Dynamik. Eine Rolling Cross-Section/Panel-Studie zu den Wirkungen des Wahlkampfes bei der vorgezogenen Bundestagswahl 2005*. Projektantrag an die Deutsche Forschungsgemeinschaft, Institut für Politikwissenschaft der Universität Duisburg-Essen.
- Schmitt-Beck, R. & Weick, S. (2001). Die dauerhafte Parteiidentifikation – nur noch ein Mythos? Eine Längsschnittanalyse zur Identifikation mit politischen Parteien in West- und Ostdeutschland. *Informationsdienst Soziale Indikatoren*, 26, 1-5.
- Schmitt-Beck, R., Weick, S. & Christoph, B. (2006). Shaky Attachments: Individual-Level Stability and Change of Partisanship among West German Voters, 1984-2001. *European Journal of Political Research*, 35 (i. E.).
- Schnell, R. (1994). *Graphisch gestützte Datenanalyse*. München: Oldenbourg.
- Schoen, H. (2005). *Wahlkampfforschung*. In J. W. Falter & H. Schoen (Eds.), *Handbuch Wahlforschung* (pp. 503-542). Wiesbaden: VS-Verlag.

- Shaw, D. R. (2000). Campaign Events, the Media and the Prospects of Victory: The 1992 and 1996 US Presidential Elections. *British Journal of Political Science*, 30, 259-289.
- Stöss, R. (1997). *Stabilität im Umbruch. Wahlbeständigkeit und Parteienwettbewerb im "Superwahljahr" 1994*. Opladen: Westdeutscher Verlag.
- Vowles, J., Aimer, P., Banducci, S. & Karp, J. (1998). *Voters' Victory? New Zealand's First Election Under Proportional Representation*. Auckland: Auckland University Press.
- Vowles, J., Aimer, P., Banducci, S., Karp, J. & Miller, R. (2004). *Voters' Veto. The 2002 Election in New Zealand and the Consolidation of Minority Government*. Auckland: Auckland University Press.
- Vowles, J., Aimer, P., Karp, J., Banducci, S., Miller, R. & Sullivan, A. (2002). *Proportional Representation on Trial*. Auckland: Auckland University Press.
- Whiteley, P. F. (1988). The Causal Relationships between Issues, Candidate Evaluations, Party Identification, and Vote Choice – the View from "Rolling Thunder". *Journal of Politics*, 50, 961-984.
- Wüst, A. (2003). Stimmung, Projektion, Prognose? In A. Wüst (Ed.), *Politbarometer* (pp. 83-107). Opladen: Leske + Budrich.
- ZDF (2006). *Politbarometer auf einen Blick*. Retrieved April 7, 2006, from <http://www.zdf.de/ZDFde/inhalt/12/0,1872,2337900,00.html>

Korrespondenzadressen

Prof. Dr. Rüdiger Schmitt-Beck
Institut für Politikwissenschaft
Universität Duisburg-Essen – Campus Duisburg
Lotharstraße 65
47057 Duisburg
E-mail: ruediger.schmitt-beck@uni-due.de

Dr. Christian Holst
Director Public Affairs /
Politik & Sozialforschung
IPSOS GmbH
Papenkamp 2-6
23879 Moelln
E-mail: Christian.Holst@Ipsos.de